# BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

## IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.**

GenCore version 4.5


OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 03:30:58 ; Search time 1643.83 Seconds
(without alignments)
166.765 Million cell updates/sec

Title: US-09-310-844B-23
Perfect score: 29
Sequence: 1 nnngauncuuunnguaagcccnangngnn 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est33:*
26: gb_est34:*
27: gb_est35:*
28: gb_est36:*
29: gb_est37:*
30: gb_est38:*
31: gb_est39:*
32: gb_est40:*
33: em_estba:*
34: em_estfun:*
35: em_esthum1:*
36: em_esthum2:*
37: em_esthum3:*
38: em_esthum4:*
39: em_esthum5:*
40: em_esthum6:*
41: em_esthum7:*
42: em_esthum8:*
43: em_esthum9:*
44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estin5:*
68: em_estom1:*
69: em_estom2:*
70: em_estov1:*
71: em_estov2:*
72: em_estpl1:*
73: em_estpl2:*
74: em_estpl3:*
75: em_estpl4:*
76: em_estpl5:*
77: em_estpl6:*
78: em_estpl7:*
79: em_estpl8:*
80: em_estpl9:*
81: em_estpl10:*
82: em_estro1:*
83: em_estro2:*
84: em_estro3:*
85: em_estro4:*
86: em_estro5:*
87: em_estro6:*
88: em_estro7:*
89: em_estro8:*
90: em_estro9:*
91: em_estro10:*
92: em_estro11:*
93: em_estro12:*
94: em_estro13:*
95: em_estro14:*
96: em_estro15:*
97: em_estro16:*
98: em_estro17:*
99: em_estro18:*
100: em_estro19:*
101: em_estro20:*
102: gb_est25:*
103: gb_est26:*
104: gb_est27:*
105: gb_est28:*
106: gb_est29:*
107: gb_est30:*
108: gb_est31:*
109: gb_est32:*
110: gb_est41:*
111: gb_est42:*
112: gb_est43:*
113: gb_est44:*
114: gb_est45:*
115: gb_est46:*
116: gb_est47:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est75:*
137: gb_est76:*
138: gb_est77:*
139: gb_est78:*
140: gb_est79:*
141: gb_est80:*
142: gb_est81:*
143: gb_est82:*
144: gb_est83:*
145: gb_est84:*
146: gb_est85:*
147: gb_est86:*
148: gb_est87:*
149: gb_est88:*
150: gb_est89:*
151: gb_est90:*
152: gb_est99:*
153: gb_est100:*
154: gb_est101:*
155: gb_est102:*
156: gb_est103:*
157: gb_est104:*
158: gb_est105:*
159: gb_est106:*
160: gb_est67:*
161: gb_est68:*
162: gb_est69:*
163: gb_est70:*
164: gb_est71:*
165: gb_est72:*
166: gb_est73:*
167: gb_est74:*
168: gb_est91:*
169: gb_est92:*
170: gb_est93:*
171: gb_est94:*
172: gb_est95:*
173: gb_est96:*
174: gb_est97:*
175: gb_est98:*
176: em_esthum29:*
177: em_esthum30:*
178: em_esthum31:*
179: em_esthum32:*
180: em_esthum33:*
181: em_estom3:*
182: em_estpl11:*
183: em_estro21:*
184: em_estro22:*
185: em_estro23:*
186: em_htc:*
187: gb_est107:*
188: gb_est108:*
189: gb_est109:*
190: gb_est110:*
191: gb_est111:*
192: gb_htc:*
193: em_gss_fun:*
194: em_gss_hum1:*
195: em_gss_hum2:*
196: em_gss_hum3:*
197: em_gss_hum4:*
198: em_gss_hum5:*
199: em_gss_hum6:*
200: em_gss_hum7:*
201: em_gss_hum8:*
202: em_gss_hum9:*
203: em_gss_inv1:*
204: em_gss_inv2:*
205: em_gss_inv3:*
206: em_gss_other:*
207: em_gss_pln1:*
208: em_gss_pln2:*
209: em_gss_pro:*
210: em_gss_rod1:*
211: em_gss_rod2:*
212: em_gss_rod3:*
213: em_gss_rod4:*
214: em_gss_rod5:*
215: em_gss_vrt1:*
216: em_gss_vrt2:*
217: em_gss_vrt3:*
218: gb_gss1:*
219: gb_gss2:*
220: gb_gss3:*
221: gb_gss4:*
222: gb_gss5:*
223: gb_gss6:*
224: gb_gss7:*
225: gb_gss8:*
226: gb_gss9:*
227: gb_gss10:*
228: gb_gss11:*
229: gb_gss12:*
230: gb_gss13:*
231: gb_gss14:*
232: gb_gss15:*
233: gb_gss16:*
234: gb_gss17:*
235: gb_gss18:*
236: gb_gss19:*
237: gb_gss20:*
238: gb_gss21:*
239: gb_gss22:*
240: gb_gss23:*
241: gb_gss24:*
242: gb_gss25:*
243: gb_gss26:*
244: gb_gss27:*
245: gb_gss28:*
246: gb_gss29:*
247: gb_gss30:*
248: gb_gss31:*
249: gb_gss32:*
250: gb_gss33:*
251: gb_gss34:*
252: em_gss_inv4:*
253: em_gss_rod6:*
254: em_gss_rod7:*
255: em_gss_rod8:*
256: gb_gss35:*
257: gb_gss36:*
258: gb_gss37:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 18 | 62.1 | 325 | 4 | AA290106 | AA290106 vb32e03.r |
| c 2 | 18 | 62.1 | 527 | 226 | AQ267968 | AQ267968 RPCI11-73 |
| c 3 | 18 | 62.1 | 596 | 226 | AQ267210 | AQ267210 RPCI11-73 |
| 4 | 17 | 58.6 | 681 | 219 | CNS004HD | AL078497 Drosophil |
| 5 | 17 | 58.6 | 943 | 219 | CNS00C4W | AL058429 Drosophil |
| 6 | 17 | 58.6 | 985 | 219 | CNS00JO5 | AL076625 Drosophil |
| 7 | 17 | 58.6 | 988 | 219 | CNS00JTC | AL076711 Drosophil |
| 8 | 17 | 58.6 | 1101 | 219 | CNS00BBM | AL056770 Drosophil |
| 9 | 17 | 58.6 | 1101 | 219 | CNS00C0H | AL058573 Drosophil |
| 10 | 16 | 55.2 | 330 | 162 | BE065429 | BE065429 RC3-BT031 |
| 11 | 16 | 55.2 | 494 | 113 | AW240011 | AW240011 ptr1c.pk0 |
| c 12 | 16 | 55.2 | 888 | 169 | BF795148 | BF795148 602256488 |
| c 13 | 15.4 | 53.1 | 145 | 150 | BF522395 | BF522395 UI-R-C3-t |
| c 14 | 15.4 | 53.1 | 272 | 26 | AV205618 | AV205618 AV205618 |
| 15 | 15.4 | 53.1 | 311 | 224 | AQ091856 | AQ091856 HS_2208_A |
| 16 | 15.4 | 53.1 | 316 | 249 | AZ762135 | AZ762135 1M0556B19 |
| 17 | 15.4 | 53.1 | 397 | 21 | AI535204 | AI535204 UI-R-C3-t |
| 18 | 15.4 | 53.1 | 408 | 224 | AQ148119 | AQ148119 HS_3109_A |
| c 19 | 15.4 | 53.1 | 417 | 223 | AQ065332 | AQ065332 HS_2227_B |
| 20 | 15.4 | 53.1 | 533 | 103 | AI914074 | AI914074 wd74e12.x |
| 21 | 15.4 | 53.1 | 554 | 164 | BE173543 | BE173543 RC2-HT056 |
| c 22 | 15.4 | 53.1 | 591 | 242 | AZ377159 | AZ377159 1M0131B22 |
| c 23 | 15.4 | 53.1 | 602 | 236 | AQ970351 | AQ970351 RPCI-23-3 |
| c 24 | 15.4 | 53.1 | 604 | 244 | AZ486340 | AZ486340 1M0314J18 |
| 25 | 15.4 | 53.1 | 613 | 149 | BF472495 | BF472495 UI-M-BH3- |
| 26 | 15.4 | 53.1 | 691 | 164 | BE214575 | BE214575 HV_CEb000 |
| c 27 | 15.4 | 53.1 | 707 | 241 | AZ291529 | AZ291529 RPCI-23-1 |
| c 28 | 15.4 | 53.1 | 828 | 220 | CNS01WBP | AL170206 Tetraodon |
| c 29 | 15.4 | 53.1 | 851 | 221 | CNS03KAZ | AL247940 Tetraodon |
| c 30 | 15.4 | 53.1 | 1048 | 221 | CNS03K91 | AL247870 Tetraodon |
| c 31 | 15.4 | 53.1 | 1086 | 221 | CNS040A8 | AL268649 Tetraodon |
| 32 | 15.4 | 53.1 | 1665 | 144 | BF120703 | BF120703 601758286 |
| 33 | 15 | 51.7 | 234 | 25 | AV130117 | AV130117 AV130117 |
| c 34 | 15 | 51.7 | 268 | 25 | AV142328 | AV142328 AV142328 |
| c 35 | 15 | 51.7 | 436 | 189 | T99738 | T99738 ye67e11.r1 |
| 36 | 15 | 51.7 | 501 | 152 | BG370061 | BG370061 LmA061 V8 |
| 37 | 15 | 51.7 | 584 | 229 | AQ465094 | AQ465094 HS_5084_B |
| c 38 | 15 | 51.7 | 660 | 230 | AQ586492 | AQ586492 RPCI-11-4 |
| c 39 | 15 | 51.7 | 915 | 153 | BG390703 | BG390703 602416413 |
| c 40 | 14.8 | 51.0 | 138 | 144 | BF077310 | BF077310 227157 MA |
| c 41 | 14.8 | 51.0 | 189 | 249 | AZ767655 | AZ767655 1M0567A06 |
| 42 | 14.8 | 51.0 | 248 | 191 | Z28552 | Z28552 HSB35G061 S |
| 43 | 14.8 | 51.0 | 289 | 25 | AV154534 | AV154534 AV154534 |
| 44 | 14.8 | 51.0 | 293 | 15 | AI057753 | AI057753 TENU1844 |
| c 45 | 14.8 | 51.0 | 316 | 147 | BF377814 | BF377814 CM2-TN013 |

ALIGNMENTS

```
RESULT   1
AA290106
LOCUS       AA290106      325 bp    mRNA            EST       14-APR-1997
DEFINITION  vb32e03.r1 Soares mouse lymph node NbMLN Mus musculus cDNA clone
            IMAGE:750652 5' similar to gb:M16762 Mouse interleukin 2 (MOUSE);,
            mRNA sequence.
ACCESSION   AA290106
VERSION     AA290106.1  GI:1936335
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 325)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:459636
            Putative full length read
            vector to vector length is 422
            Seq primer: -28m13 rev2 ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .325
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:750652"
                     /clone_lib="Soares mouse lymph node NbMLN"
                     /sex="male"
                     /tissue_type="lymph node"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
                     /note="Organ: lymph node; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Not I - oligo(dT) primer
                     [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGATACTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. RNA
                     provided by Dr. Bertrand Jordan. Library constructed and
                     normalized by Bento Soares and M.Fatima Bonaldo."
BASE COUNT      107 a     47 c     41 g    130 t
ORIGIN


  Query Match             62.1%;  Score 18;  DB 4;  Length 325;
  Best Local Similarity   54.2%;  Pred. No. 9.5;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       4 gauncuuunnguaagcccnangng 27
           ||: |::: |:|||||| | | |
Db     170 GATTCTTTTGTAAGCCCCAAGGG 193


RESULT   2
AQ267968/c
LOCUS       AQ267968      527 bp    DNA             GSS       27-APR-1999
DEFINITION  RPCI11-73E11.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-73E11,
            DNA sequence.
ACCESSION   AQ267968
VERSION     AQ267968.1  GI:3795572
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 527)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
  TITLE     Use of human BAC End Sequences for Sequence-Ready Map Building
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: RPCI11-73E11.TK
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
```

```
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: SP6
            Class: BAC ends.
FEATURES              Location/Qualifiers
     source           1. .527
                      /organism="Homo sapiens"
                      /db_xref="GDB:7527754"
                      /db_xref="taxon:9606"
                      /clone="RPCI-11-73E11"
                      /clone_lib="RPCI-11"
                      /sex="Male"
                      /cell_type="Lymphocytes"
                      /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                      RPCI11 Human Male BAC Library"
BASE COUNT       203 a      78 c      77 g     167 t       2 others
ORIGIN


  Query Match             62.1%;  Score 18;  DB 226;  Length 527;
  Best Local Similarity   54.2%;  Pred. No. 11;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       4 gauncuuunnguaagcccnangng 27
           ||: |:::  |:|||||| | | |
Db     355 GATTCTTTTTGTAAGCCCTAGGGG 332


RESULT   3
AQ267210/c
LOCUS       AQ267210       596 bp    DNA             GSS       27-APR-1999
DEFINITION  RPCI11-73C7.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-73C7, DNA
            sequence.
ACCESSION   AQ267210
VERSION     AQ267210.1  GI:3794814
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE 1  (bases 1 to 596)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
  TITLE     Use of human BAC End Sequences for Sequence-Ready Map Building
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: RPCI11-73C7.TK
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: SP6
            Class: BAC ends.
FEATURES              Location/Qualifiers
     source           1. .596
                      /organism="Homo sapiens"
                      /db_xref="GDB:7527702"
                      /db_xref="taxon:9606"
                      /clone="RPCI-11-73C7"
                      /clone_lib="RPCI-11"
                      /sex="Male"
                      /cell_type="Lymphocytes"
                      /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                      RPCI11 Human Male BAC Library"
BASE COUNT       227 a      92 c      88 g     186 t       3 others
ORIGIN


  Query Match             62.1%;  Score 18;  DB 226;  Length 596;
  Best Local Similarity   54.2%;  Pred. No. 11;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       4 gauncuuunnguaagcccnangng 27
           ||: |:::  |:|||||| | | |
Db     355 GATTCTTTTTGTAAGCCCTAGGGG 332


RESULT   4
CNS004HD
LOCUS       CNS004HD       681 bp    DNA             GSS       03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence T7 end of BAC #
            BACR10O24 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL078497
VERSION     AL078497.1  GI:4972152
KEYWORDS    GSS.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
  REFERENCE 1  (bases 1 to 681)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES              Location/Qualifiers
     source           1. .681
                      /organism="Drosophila melanogaster"
                      /db_xref="taxon:7227"
                      /clone_lib="RPCI-98"
                      /clone="BACR10O24"
                      /note="end : T7"
BASE COUNT       227 a     126 c     111 g     180 t      37 others
ORIGIN


  Query Match             58.6%;  Score 17;  DB 219;  Length 681;
  Best Local Similarity   56.5%;  Pred. No. 42;
  Matches   13;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       5 auncuuunnguaagcccnangng 27
           |: |:::  |:|||||| | |||
Db     293 ATCCTTTGGGTAAGCCCGAAGNG 315
```

```
RESULT   5
CNS00C4W
LOCUS       CNS00C4W      943 bp    DNA             GSS       04-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence T7 end of BAC #
            BACR25E15 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL058429
VERSION     AL058429.1  GI:4946310
KEYWORDS    GSS.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 943)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
  MMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES              Location/Qualifiers
     source           1. .943
                      /organism="Drosophila melanogaster"
                      /db_xref="taxon:7227"
                      /clone_lib="RPCI-98"
                      /clone="BACR25E15"
                      /note="end : T7"
BASE COUNT      331 a    166 c    143 g    265 t     38 others
ORIGIN


  Query Match             58.6%;  Score 17;  DB 219;  Length 943;
  Best Local Similarity   52.2%;  Pred. No. 45;
  Matches   12;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       5 auncuuunnguaagcccnangng 27
           |: |:::  |:|||||| | | |
Db     398 ATCCTTTGGGTAAGCCCGAAGTG 420


RESULT   6
CNS00JO5
LOCUS       CNS00JO5      985 bp    DNA             GSS       03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC:
            BACR38L15 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL076625
VERSION     AL076625.1  GI:4956102
KEYWORDS    GSS.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 985)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES              Location/Qualifiers
     source           1. .985
                      /organism="Drosophila melanogaster"
                      /db_xref="taxon:7227"
                      /clone_lib="RPCI-98"
                      /clone="BACR38L15"
                      /note="end : TET3"
BASE COUNT      323 a    172 c    148 g    303 t     39 others
ORIGIN


  Query Match             58.6%;  Score 17;  DB 219;  Length 985;
  Best Local Similarity   52.2%;  Pred. No. 45;
  Matches   12;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       5 auncuuunnguaagcccnangng 27
           |: |:::  |:|||||| | | |
Db     462 ATCCTTTGGGTAAGCCCGAAGTG 484


RESULT   7
CNS00JTC
LOCUS       CNS00JTC      988 bp    DNA             GSS       03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC:
            BACR39I07 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL076711
VERSION     AL076711.1  GI:4956289
KEYWORDS    GSS.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 988)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
```

```
FEATURES             Location/Qualifiers
     source          1. .988
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone_lib="RPCI-98"
                     /clone="BACR39I07"
                     /note="end : TET3"
BASE COUNT      322 a    171 c    151 g    313 t     31 others
ORIGIN


  Query Match             58.6%;  Score 17;  DB 219;  Length 988;
  Best Local Similarity   52.2%;  Pred. No. 45;
  Matches   12;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      5 auncuuunnguaagcccnangng 27
          |: |:::  |:|||||| | | |
Db    448 ATCCTTTGGGTAAGCCCGAAGTG 470


ULT   8
NS00BBM
LOCUS       CNS00BBM     1101 bp    DNA             GSS       04-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR22G16 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL056770
VERSION     AL056770.1  GI:4937438
KEYWORDS    GSS.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .1101
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone_lib="RPCI-98"
                     /clone="BACR22G16"
                     /note="end : TET3"
BASE COUNT      368 a    203 c    165 g    333 t     32 others
ORIGIN


  Query Match             58.6%;  Score 17;  DB 219;  Length 1101;
  Best Local Similarity   52.2%;  Pred. No. 46;
  Matches   12;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      5 auncuuunnguaagcccnangng 27
          |: |:::  |:|||||| | | |
Db    454 ATCCTTTGGGTAAGCCCGAAGTG 476


RESULT   9
CNS00C0H
LOCUS       CNS00C0H     1101 bp    DNA             GSS       04-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR24D19 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL058573
VERSION     AL058573.1  GI:4946151
KEYWORDS    GSS.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .1101
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone_lib="RPCI-98"
                     /clone="BACR24D19"
                     /note="end : TET3"
BASE COUNT      356 a    196 c    163 g    317 t     69 others
ORIGIN


  Query Match             58.6%;  Score 17;  DB 219;  Length 1101;
  Best Local Similarity   52.2%;  Pred. No. 46;
  Matches   12;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      5 auncuuunnguaagcccnangng 27
          |: |:::  |:|||||| | | |
Db    473 ATCCTTTGGGTAAGCCCGAAGTG 495


RESULT  10
BE065429
LOCUS       BE065429     330 bp    mRNA            EST       09-JUN-2000
DEFINITION  RC3-BT0316-071299-011-e01 BT0316 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE065429
VERSION     BE065429.1  GI:8410079
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 330)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
```

```
          ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
          Simpson,A.J.
  TITLE   Shotgun sequencing of the human transcriptome with ORF expressed
          sequence tags
  JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE 20202663
COMMENT   Contact: Simpson A.J.G.
          Laboratory of Cancer Genetics
          Ludwig Institute for Cancer Research
          Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
          Brazil
          Tel: +55-11-2704922
          Fax: +55-11-2707001
          Email: asimpson@ludwig.org.br
          This sequence was derived from the FAPESP/LICR Human Cancer Genome
          Project. This entry can be seen in the following URL
          (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=RC3-BT0316-071
          299-011-e01&t3=1999-12-07&t4=1)
          Seq primer: puc 18 forward
          High quality sequence stop: 330.
  ATURES            Location/Qualifiers
     source         1. .330
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone_lib="BT0316"
                    /dev_stage="Adult"
                    /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
                    SmaI; A mini-library was made by cloning products derived
                    from  ORESTES PCR (U.S. Letters Patent application No. 196
                    ,716 -  Ludwig Institute for Cancer Research) profiles
                    into  the pUC 18 vector. Reverse transcription of tissue
                    mRNA and cDNA amplification were performed under low
                    stringency conditions."
BASE COUNT      100 a     64 c     73 g     93 t
ORIGIN


  Query Match             55.2%;  Score 16;  DB 162;  Length 330;
  Best Local Similarity   52.4%;  Pred. No. 1.3e+02;
  Matches   11;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      5 auncuuunnguaagcccnang 25
          |: |:::  |:|||||| | |
Db    184 ATGCTTTATGTAAGCCCAAGG 204


RESULT  11
AW240011
LOCUS       AW240011      494 bp    mRNA            EST       27-APR-2000
DEFINITION  ptr1c.pk001.120 chicken T cell cDNA library Gallus gallus cDNA
          clone ptr1c.pk001.120 5', mRNA sequence.
ACCESSION   AW240011
VERSION     AW240011.1  GI:6579751
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
          Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 494)
  AUTHORS   Morgan,R.
  TITLE     Chicken T cell ESTs
  JOURNAL   Unpublished (1999)
COMMENT   Contact: Robin Morgan
          University of Delaware
          Townsend Hall, Newark, DE 19717, USA
          Tel: 302-831-1341
          Fax: 302-831-2822
          Email: morgan@udel.edu
          Clones can be ordered online at http://www.chickest.udel.edu.
FEATURES            Location/Qualifiers
     source         1. .494
                    /organism="Gallus gallus"
                    /db_xref="taxon:9031"
                    /clone="ptr1c.pk001.120"
                    /clone_lib="chicken T cell cDNA library"
                    /sex="Male and Female"
                    /cell_type="splenic T cell"
                    /lab_host="E.coli DH10B"
                    /note="Vector: pB42AD"
BASE COUNT      151 a     82 c     96 g    143 t     22 others
ORIGIN


  Query Match             55.2%;  Score 16;  DB 113;  Length 494;
  Best Local Similarity   60.0%;  Pred. No. 1.4e+02;
  Matches   12;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      4 gauncuuunnguaagcccna 23
          ||: |:::  |:||||||||
Db     18 GATGCTTTTTGTAAGCCCNA 37


RESULT  12
BF795148/c
LOCUS       BF795148      888 bp    mRNA            EST       12-JAN-2001
DEFINITION  602256488F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4339844 5',
          mRNA sequence.
ACCESSION   BF795148
VERSION     BF795148.1  GI:12100202
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 888)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: Louis Staudt, M.D., Ph.D.
           cDNA Library Preparation: Life Technologies, Inc.
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Incyte Genomics, Inc.
           Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM9950  row: p  column: 21
          High quality sequence stop: 678.
FEATURES            Location/Qualifiers
     source         1. .888
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:4339844"
                    /clone_lib="NIH_MGC_85"
                    /tissue_type="lymphoma, cell line"
                    /lab_host="DH10B (phage-resistant)"
                    /note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
                    Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                    Average insert size 1.867 kb. Library enriched for
                    full-length clones and constructed by Life Technologies.
                    Note: this is a NIH_MGC Library."
BASE COUNT      217 a    217 c    236 g    218 t
ORIGIN


  Query Match             55.2%;  Score 16;  DB 169;  Length 888;
  Best Local Similarity   50.0%;  Pred. No. 1.7e+02;
  Matches   11;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      6 uncuuunnguaagcccnangng 27
          : |:::  |:|||||| | | |
Db    817 TCCTTTCCGTAAGCCCCAAGTG 796
```

```
RESULT 13
BF522395/c
LOCUS       BF522395      145 bp    mRNA            EST       11-DEC-2000
DEFINITION  UI-R-C3-ta-e-05-0-UI.r2 UI-R-C3 Rattus norvegicus cDNA clone
            UI-R-C3-ta-e-05-0-UI 5', mRNA sequence.
ACCESSION   BF522395
VERSION     BF522395.1  GI:11630362
KEYWORDS    EST.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 145)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
  MMENT     Contact: Soares, MB
            Program for Rat Gene Discovery and Mapping
            University of Iowa
            451 Eckstein Medical Research Building Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: msoares@blue.weeg.uiowa.edu
            cDNA Library Preparation: M.B. Soares Lab Clone distribution:
            clones will be available through Research Genetics (www.resgen.com)
            This clone is also available through the I.M.A.G.E. Consortium at
            LLNL (info@image.llnl.gov). IMAGE ID= 1769254
            Seq primer: M13 Forward.
FEATURES             Location/Qualifiers
     source          1. .145
                     /organism="Rattus norvegicus"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /clone="UI-R-C3-ta-e-05-0-UI"
                     /clone_lib="UI-R-C3"
                     /dev_stage="adult"
                     /lab_host="DH10B (Life Technologies)"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-C3
                     library is a subtracted library of a series, ultimately
                     derived from a mixture of individually tagged normalized
                     libraries from rat placenta, adult lung, brain, liver,
                     kidney, heart, spleen, ovary, muscle, and 8, 12 and 18-day
                     embryos, after a series of subtractions to reduce the
                     representation of cDNAs from which ESTs had already been
                     generated. The following serially subtracted libraries
                     were generated in this process: UI-R-C3, UI-R-C2p, UI-R-C1
                     , UI-R-C0, UI-R-A1, UI-R-E1. The tag is a string of 3-5
                     nucleotides present between the Not I site and the
                     oligo-dT track whichallows identification of the library
                     of origin of a clone within themixture. The subtracted
                     library (UI-R-C3) was constructed as follows: PCRamplified
                     cDNA inserts from UI-R-C2p clones from which 3' ESTs had
                     been derived was used as a driver in a hybridization with
                     the UI-R-C2p library in the form of single-stranded
                     circles. The remaining single-stranded circles (subtracted
                     library) was purified by hydroxyapatite column
                     chromatography, converted to double-stranded circles and
                     electroporated into DH10B bacteria (LifeTechnologies) to
                     generate the UI-R-C3 library. This procedure has been
                     previously described (Bonaldo, Lennon and Soares, Genome
                     Research 6:791-806, 1996)"
BASE COUNT       23 a     39 c     40 g     43 t
ORIGIN
```

```
  Query Match             53.1%;  Score 15.4;  DB 150;  Length 145;
  Best Local Similarity   50.0%;  Pred. No. 2.4e+02;
  Matches   11;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       4 gauncuuunnguaagcccnang 25
           ||: |::: :||||| | |
Db      75 GATGCTTTTCCTAAGCCCCAGG 54
```

```
RESULT 14
AV205618/c
LOCUS       AV205618      272 bp    mRNA            EST       30-OCT-1999
DEFINITION  AV205618 RIKEN full-length enriched, adult male testis Mus musculus
            cDNA clone 1700081C10 3', mRNA sequence.
ACCESSION   AV205618
VERSION     AV205618.1  GI:6146471
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 272)
  AUTHORS   Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Carninci,P., Endo,T.,
            Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F.,
            Ishii,Y., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai
            ,C., Kawai,J., Kikuchi,N., Kojima,Y., Koya,S., Kusakabe,M.,
            Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.,
            Owa,C., Ozawa,Y., Saito,H., Sano,M., Sato,K., Shibata,K., Shibata
            ,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H.,
            Suzuki,H., Takahashi,F., Tateno,M., Tominaga,N., Tsunoda,Y.,
            Watahiki,A., Watanabe,S., Yamamura,T., Yasunishi,A., Yokota,T.,
            Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Konno,H., et al. 1999)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Yoshihide Hayashizaki
            Genome Exploration Research Group, Life Science Tsukuba Center,
            Genome Science Laboratory
            The Institute of Physical and Chemical Research (RIKEN), Genomic
            Sciences Center
            3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
            Tel: +81-298-36-9013
            Fax: +81-298-36-9098
            Email: genome-res@rtc.riken.go.jp,
            URL:http://genome.rtc.riken.go.jp/
            Sasaki,N., Izawa,M., Watahiki,M., Ozawa,K., Tanaka,T., Yoneda,Y.,
            Matsuura,S., Carninci,P., Muramatsu,M., Okazaki,Y. and Hayashizaki
            ,Y.
             Transcriptional sequencing: A method for DNA sequencing using RNA
            polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)
             Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
            Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki
            ,Y. and Hayashizaki,Y.
             Automated filtration-based high-throughput plasmid preparation
            system. Genome Res. 9 (5), 463-470 (1999)
             Carninci,P. and Hayashizaki,Y.
             High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
            19-44 (1999)
             Please visit our web site (http://genome.rtc.riken.go.jp) for
            further details.
FEATURES             Location/Qualifiers
     source          1. .272
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="1700081C10"
                     /clone_lib="RIKEN full-length enriched, adult male testis"
                     /sex="male"
                     /tissue_type="testis"
                     /dev_stage="adult"
                     /lab_host="SOLR"
                     /note="Site_1: XhoI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
```

GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGAGCGGCCGCAATTAATTCTCGAGTTAATTAAATTAATCCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. "
BASE COUNT 90 a 52 c 69 g 61 t
ORIGIN

Query Match 53.1%; Score 15.4; DB 26; Length 272;
Best Local Similarity 50.0%; Pred. No. 2.8e+02;
Matches 11; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 4 gauncuuunnguaagcccnang 25
||: |::: |:|||| | | |
Db 112 GATGCTTTCTGTAAGCTCCAGG 91

SULT 15
091856
LOCUS AQ091856 311 bp DNA GSS 26-AUG-1998
DEFINITION HS_2208_A1_H11_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2208 Col=21 Row=O, DNA sequence.
ACCESSION AQ091856
VERSION AQ091856.1 GI:3460767
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 311)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999).
MEDLINE 99380589
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 2208 row: O column: 21
Class: BAC ends
High quality sequence stop: 311.
FEATURES Location/Qualifiers
source 1. .311
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=2208 Col=21 Row=O"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in E-Coli DH10B"
BASE COUNT 134 a 28 c 51 g 95 t 3 others
ORIGIN

Query Match 53.1%; Score 15.4; DB 224; Length 311;
Best Local Similarity 50.0%; Pred. No. 2.9e+02;
Matches 11; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 4 gauncuuunnguaagcccnang 25
||: ::: |:|||||| | |
Db 217 GATGTTTTATGTAAGCCCCATG 238

Search completed: October 2, 2001, 04:38:38
Job time: 4060 sec

GenCore version 4.5


OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 03:33:13 ; Search time 57.41 Seconds
(without alignments)
95.628 Million cell updates/sec

Title: US-09-310-844B-23
Perfect score: 29
Sequence: 1 nnngauncuuunnguaagcccnangngnn 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 94655562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| | Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|---|
| | 1 | 18 | 62.1 | 209 | 1 | US-08-510-032A-8 | Sequence 8, Appli |
| | 2 | 18 | 62.1 | 209 | 3 | US-08-688-514-8 | Sequence 8, Appli |
| | 3 | 18 | 62.1 | 801 | 6 | 5314995-8 | Patent No. 5314995 |
| : | 4 | 18 | 62.1 | 8491 | 2 | US-08-757-439-1 | Sequence 1, Appli |
| c | 5 | 14.4 | 49.7 | 1894 | 4 | US-09-004-731-29 | Sequence 29, Appl |
| | 6 | 14.4 | 49.7 | 1894 | 4 | US-09-004-731-31 | Sequence 31, Appl |
| c | 7 | 14.4 | 49.7 | 1894 | 4 | US-09-032-215-3 | Sequence 3, Appli |
| c | 8 | 14.4 | 49.7 | 1894 | 4 | US-08-749-699-29 | Sequence 29, Appl |
| | 9 | 14.4 | 49.7 | 1894 | 4 | US-08-749-699-31 | Sequence 31, Appl |
| | 10 | 14.4 | 49.7 | 10409 | 3 | US-08-772-440-33 | Sequence 33, Appl |
| | 11 | 14.4 | 49.7 | 21126 | 1 | US-08-008-216-19 | Sequence 19, Appl |
| | 12 | 14.4 | 49.7 | 21126 | 1 | US-08-459-569-19 | Sequence 19, Appl |
| | 13 | 14.4 | 49.7 | 21126 | 1 | US-08-458-831-19 | Sequence 19, Appl |
| | 14 | 13.8 | 47.6 | 2670 | 1 | US-08-121-713D-61 | Sequence 61, Appl |
| | 15 | 13.8 | 47.6 | 2670 | 1 | US-08-835-268-61 | Sequence 61, Appl |
| | 16 | 13.8 | 47.6 | 2670 | 2 | US-09-060-692-61 | Sequence 61, Appl |
| | 17 | 13.8 | 47.6 | 2670 | 3 | US-08-833-391-61 | Sequence 61, Appl |
| | 18 | 13.8 | 47.6 | 2670 | 5 | PCT-US94-10151A-61 | Sequence 61, Appl |
| | 19 | 13.4 | 46.2 | 2829 | 3 | US-08-851-843A-53 | Sequence 53, Appl |
| | 20 | 13.4 | 46.2 | 2829 | 4 | US-08-974-549A-220 | Sequence 220, App |
| c | 21 | 13.2 | 45.5 | 945 | 4 | US-09-434-774-15 | Sequence 15, Appl |
| c | 22 | 13.2 | 45.5 | 3867 | 3 | US-08-762-428A-5 | Sequence 5, Appli |
| c | 23 | 13.2 | 45.5 | 4016 | 3 | US-08-762-428A-7 | Sequence 7, Appli |
| | 24 | 13.2 | 45.5 | 4884 | 1 | US-07-665-792E-10 | Sequence 10, Appl |
| c | 25 | 13 | 44.8 | 172 | 1 | US-08-480-910-8 | Sequence 8, Appli |
| c | 26 | 13 | 44.8 | 300 | 1 | US-08-480-910-2 | Sequence 2, Appli |
| c | 27 | 13 | 44.8 | 1285 | 1 | US-08-480-910-1 | Sequence 1, Appli |
| c | 28 | 13 | 44.8 | 1285 | 5 | PCT-US95-00052-1 | Sequence 1, Appli |
| | 29 | 12.8 | 44.1 | 423 | 1 | US-08-470-179-41 | Sequence 41, Appl |
| | 30 | 12.8 | 44.1 | 423 | 1 | US-08-470-179-120 | Sequence 120, App |
| | 31 | 12.8 | 44.1 | 423 | 1 | US-08-470-179-123 | Sequence 123, App |
| | 32 | 12.8 | 44.1 | 705 | 4 | US-08-998-416-1136 | Sequence 1136, Ap |
| | 33 | 12.8 | 44.1 | 1210 | 3 | US-09-002-298-2 | Sequence 2, Appli |
| | 34 | 12.8 | 44.1 | 1212 | 3 | US-09-188-930-249 | Sequence 249, App |
| | 35 | 12.8 | 44.1 | 1803 | 3 | US-08-458-922-2 | Sequence 2, Appli |
| | 36 | 12.8 | 44.1 | 2221 | 1 | US-08-418-782-2 | Sequence 2, Appli |
| | 37 | 12.8 | 44.1 | 2221 | 1 | US-08-228-662-2 | Sequence 2, Appli |
| | 38 | 12.8 | 44.1 | 2221 | 2 | US-08-852-219-2 | Sequence 2, Appli |
| | 39 | 12.8 | 44.1 | 2235 | 1 | US-08-418-782-1 | Sequence 1, Appli |
| | 40 | 12.8 | 44.1 | 2235 | 1 | US-08-228-662-1 | Sequence 1, Appli |
| | 41 | 12.8 | 44.1 | 2235 | 2 | US-08-852-219-1 | Sequence 1, Appli |
| | 42 | 12.8 | 44.1 | 2331 | 1 | US-08-418-782-20 | Sequence 20, Appl |
| | 43 | 12.8 | 44.1 | 2331 | 2 | US-08-852-219-20 | Sequence 20, Appl |
| | 44 | 12.8 | 44.1 | 2955 | 2 | US-08-867-941-4 | Sequence 4, Appli |
| | 45 | 12.8 | 44.1 | 2955 | 4 | US-09-074-658-4 | Sequence 4, Appli |

ALIGNMENTS

```
RESULT   1
US-08-510-032A-8
; Sequence 8, Application US/08510032A
; Patent No. 5712126
;  GENERAL INFORMATION:
;    APPLICANT:  Sherman Weissman and Yarindra Prashar
;    TITLE OF INVENTION:  Analysis of Gene Expression By Display of 3'-
;    TITLE OF INVENTION:  end Restriction Fragments of cDNA
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Yahwak & Associates
;      STREET:  25 Skytop Drive
;      CITY:  Trumbull
;      STATE:  Connecticut
;      COUNTRY:  USA
;      ZIP:  06611
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  floppy disk
;      COMPUTER:  Macintosh
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  Microsoft Word 4.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/510,032A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  George M. Yahwak
;      REGISTRATION NUMBER:  26,824
;      REFERENCE/DOCKET NUMBER:  Yale
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (203)268-1951
;      TELEFAX:  (203)268-1951
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  209 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-08-510-032A-8

  Query Match             62.1%;  Score 18;  DB 1;  Length 209;
  Best Local Similarity   54.2%;  Pred. No. 0.064;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       4 gauncuuunnguaagcccnangng 27
            ||: |:::  |:|||||| | | |
Db      36 GATTCTTTTTGTAAGCCCTAGGGG 59
```

```
RESULT   2
US-08-688-514-8
; Sequence 8, Application US/08688514
; Patent No. 6010850
;  GENERAL INFORMATION:
;    APPLICANT:  Sherman Weissman and Yarindra Prashar
;    TITLE OF INVENTION:  Analysis of Gene Expression By Display of 3'-
;    TITLE OF INVENTION:  end Restriction Fragments of cDNA
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Yahwak & Associates
;      STREET:  25 Skytop Drive
;      CITY:  Trumbull
;      STATE:  Connecticut
;      COUNTRY:  USA
;      ZIP:  06611
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  floppy disk
       COMPUTER:  Macintosh
       OPERATING SYSTEM:  MS-DOS
       SOFTWARE:  Microsoft Word 4.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/688,514
;      FILING DATE:
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  George M. Yahwak
;      REGISTRATION NUMBER:  26,824
;      REFERENCE/DOCKET NUMBER:  Yale
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (203)268-1951
;      TELEFAX:  (203)268-1951
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  209 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-08-688-514-8


  Query Match             62.1%;  Score 18;  DB 3;  Length 209;
  Best Local Similarity   54.2%;  Pred. No. 0.064;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       4 gauncuuunnguaagcccnangng 27
           ||: |:::  |:||||| | | |
Db      36 GATTCTTTTTGTAAGCCCTAGGGG 59


RESULT   3
5314995-8
;Patent No. 5314995
;    APPLICANT: FELL, HENRY P.;GAYLE, MARGIT A.
;    TITLE OF INVENTION: THERAPEUTIC INTERLEUKIN-2-ANTIBODY
;BASED FUSION PROTEINS
;    NUMBER OF SEQUENCES: 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/468,390
;      FILING DATE: 22-JAN-1990
;SEQ ID NO:8:
;      LENGTH: 801
5314995-8


  Query Match             62.1%;  Score 18;  DB 6;  Length 801;
  Best Local Similarity   54.2%;  Pred. No. 0.086;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       4 gauncuuunnguaagcccnangng 27
           ||: |:::  |:||||| | | |
Db     653 gattctttttgtaagccctagggg 676


RESULT   4
US-08-757-439-1/c
; Sequence 1, Application US/08757439
; Patent No. 5866371
;  GENERAL INFORMATION:
;    APPLICANT:  BADZIONG, Werner
;    APPLICANT:  HABERMANN, Paul
;    APPLICANT:  MOELLER, Joerg
;    APPLICANT:  ARETZ, Werner
;    TITLE OF INVENTION:  PROCESS FOR USING THE YEAST ADH II
;    TITLE OF INVENTION:  PROMOTER SYSTEM FOR THE PRODUCTION OF HETEROLOGOUS
;    TITLE OF INVENTION:  PROTEINS IN HIGH YIELDS
;    NUMBER OF SEQUENCES:  1
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/757,439
;      FILING DATE:  27-NOV-1996
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE 19544233.4
;      FILING DATE:  28-NOV-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  SANDERCOCK, Colin G.
;      REGISTRATION NUMBER:  31,298
;      REFERENCE/DOCKET NUMBER:  18748/303/HOCE
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8491 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-757-439-1


  Query Match             62.1%;  Score 18;  DB 2;  Length 8491;
  Best Local Similarity   54.2%;  Pred. No. 0.15;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       4 gauncuuunnguaagcccnangng 27
           ||: |:::  |:||||| | | |
Db    6651 GATTCTTTTTGTAAGCCCTAGGGG 6628


RESULT   5
US-09-004-731-29/c
; Sequence 29, Application US/09004731
; Patent No. 6177258
;  GENERAL INFORMATION:
;    APPLICANT:  Wu Hunter, Shirley
;    APPLICANT:  Stiegler, Gary
;    APPLICANT:  Gaines, Patrick J.
;    TITLE OF INVENTION:  FLEA PROTEASE PROTEINS, NUCLEIC ACID
;    TITLE OF INVENTION:  MOLECULES AND USES THEREOF
;    NUMBER OF SEQUENCES:  103
;    CORRESPONDENCE ADDRESS:
```

```
;       ADDRESSEE:  Sheridan Ross P.C.
;       STREET:  1700 Lincoln Street, Suite 3500
;       CITY:  Denver
;       STATE:  Colorado
;       COUNTRY:  USA
;       ZIP:  80203
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/004,731
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/749,699
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Connell, Gary J.
;       REGISTRATION NUMBER:  32,020
;       REFERENCE/DOCKET NUMBER:  2618-25-C3
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (303) 863-9700
;       TELEFAX:  (303) 863-0223
;   INFORMATION FOR SEQ ID NO:  29:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1894 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  335..1535
US-09-004-731-29


  Query Match             49.7%;  Score 14.4;  DB 4;  Length 1894;
  Best Local Similarity   47.6%;  Pred. No. 15;
  Matches   10;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       5 auncuuunnguaagcccnang 25
           |: |:::  |:|||| | | |
Db    1567 ATACTTTGGGTAAGCTCGATG 1547


RESULT   6
US-09-004-731-31
; Sequence 31, Application US/09004731
; Patent No. 6177258
;  GENERAL INFORMATION:
;    APPLICANT:  Wu Hunter, Shirley
;    APPLICANT:  Stiegler, Gary
;    APPLICANT:  Gaines, Patrick J.
;    TITLE OF INVENTION:  FLEA PROTEASE PROTEINS, NUCLEIC ACID
;    TITLE OF INVENTION:  MOLECULES AND USES THEREOF
;    NUMBER OF SEQUENCES:  103
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sheridan Ross P.C.
;      STREET:  1700 Lincoln Street, Suite 3500
;      CITY:  Denver
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80203
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/004,731
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/749,699
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Connell, Gary J.
;      REGISTRATION NUMBER:  32,020
;      REFERENCE/DOCKET NUMBER:  2618-25-C3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 863-9700
;      TELEFAX:  (303) 863-0223
;  INFORMATION FOR SEQ ID NO:  31:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1894 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-09-004-731-31


  Query Match             49.7%;  Score 14.4;  DB 4;  Length 1894;
  Best Local Similarity   47.6%;  Pred. No. 15;
  Matches   10;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       5 auncuuunnguaagcccnang 25
           |: |:::  |:|||| | | |
Db    328 ATACTTTGGGTAAGCTCGATG 348


RESULT   7
US-09-032-215-3/c
; Sequence 3, Application US/09032215
; Patent No. 6204010
;  GENERAL INFORMATION:
;    APPLICANT:  Stiegler, Gary L.
;    APPLICANT:  Gaines, Patrick J.
;    TITLE OF INVENTION:  FLEA PROTEASE PROTEINS, NUCLEIC
;    TITLE OF INVENTION:  ACID MOLECULES, AND USES THEREOF
;    NUMBER OF SEQUENCES:  50
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sheridan Ross P.C.
;      STREET:  1700 Lincoln Street, Suite 3500
;      CITY:  Denver
;      STATE:  Colorado
;      COUNTRY:  U.S.A.
;      ZIP:  80203
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII DOS TEXT
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/032,215
;      FILING DATE:  27-FEB-1998
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Connell, Gary J.
;      REGISTRATION NUMBER:  32,020
;      REFERENCE/DOCKET NUMBER:  2618-25-C6
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 863-9700
;      TELEFAX:  (303) 863-0223
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1894 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
```

```
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  335..1534
US-09-032-215-3


  Query Match             49.7%;  Score 14.4;  DB 4;  Length 1894;
  Best Local Similarity   47.6%;  Pred. No. 15;
  Matches   10;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      5 auncuuunnguaagcccnang 25
          |: |:::  |:|||| | | |
Db   1567 ATACTTTGGGTAAGCTCGATG 1547


RESULT   8
US-08-749-699-29/c
; Sequence 29, Application US/08749699
; Patent No. 6210920
   GENERAL INFORMATION:
     APPLICANT:  Wu Hunter, Shirley
     APPLICANT:  Stiegler, Gary
;    APPLICANT:  Gaines, Patrick J.
;    TITLE OF INVENTION:  FLEA PROTEASE PROTEINS, NUCLEIC ACID
;    TITLE OF INVENTION:  MOLECULES AND USES THEREOF
;    NUMBER OF SEQUENCES:  103
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sheridan Ross P.C.
;      STREET:  1700 Lincoln Street, Suite 3500
;      CITY:  Denver
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80203
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/749,699
;      FILING DATE:
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Connell, Gary J.
;      REGISTRATION NUMBER:  32,020
;      REFERENCE/DOCKET NUMBER:  2618-25-C3
;    TELECOMMUNICATION INFORMATION:
       TELEPHONE:  (303) 863-9700
       TELEFAX:  (303) 863-0223
   INFORMATION FOR SEQ ID NO:  29:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1894 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  335..1535
US-08-749-699-29


  Query Match             49.7%;  Score 14.4;  DB 4;  Length 1894;
  Best Local Similarity   47.6%;  Pred. No. 15;
  Matches   10;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      5 auncuuunnguaagcccnang 25
          |: |:::  |:|||| | | |
Db   1567 ATACTTTGGGTAAGCTCGATG 1547


RESULT   9
US-08-749-699-31
; Sequence 31, Application US/08749699
; Patent No. 6210920
;  GENERAL INFORMATION:
;    APPLICANT:  Wu Hunter, Shirley
;    APPLICANT:  Stiegler, Gary
;    APPLICANT:  Gaines, Patrick J.
;    TITLE OF INVENTION:  FLEA PROTEASE PROTEINS, NUCLEIC ACID
;    TITLE OF INVENTION:  MOLECULES AND USES THEREOF
;    NUMBER OF SEQUENCES:  103
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sheridan Ross P.C.
;      STREET:  1700 Lincoln Street, Suite 3500
;      CITY:  Denver
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80203
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/749,699
;      FILING DATE:
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Connell, Gary J.
;      REGISTRATION NUMBER:  32,020
;      REFERENCE/DOCKET NUMBER:  2618-25-C3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 863-9700
;      TELEFAX:  (303) 863-0223
;  INFORMATION FOR SEQ ID NO:  31:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1894 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-749-699-31


  Query Match             49.7%;  Score 14.4;  DB 4;  Length 1894;
  Best Local Similarity   47.6%;  Pred. No. 15;
  Matches   10;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      5 auncuuunnguaagcccnang 25
          |: |:::  |:|||| | | |
Db    328 ATACTTTGGGTAAGCTCGATG 348


RESULT  10
US-08-772-440-33
; Sequence 33, Application US/08772440
; Patent No. 6046158
;  GENERAL INFORMATION:
;    APPLICANT:  Ariizumi, Kiyoshi
;    APPLICANT:  Takashima, Akira
;    TITLE OF INVENTION:  UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE
;    TITLE OF INVENTION:  LECTINS, DECTIN-1 AND DECTIN-2; COMPOSITIONS AND USES
;    TITLE OF INVENTION:  THEREOF
;    NUMBER OF SEQUENCES:  42
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
```

```
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/772,440
;       FILING DATE:  CONCURRENTLY HEREWITH
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Parker, David L.
;       REGISTRATION NUMBER:  32,165
;       REFERENCE/DOCKET NUMBER:  UTXD:493
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  512/418-3000
;       TELEFAX:  512/474-7577
;   INFORMATION FOR SEQ ID NO:  33:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  10409 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     FEATURE:
        NAME/KEY:  modified_base
        LOCATION:  6510
        OTHER INFORMATION:  /mod_base= OTHER
;       OTHER INFORMATION:  /note= "D = A or G or T"
;     FEATURE:
;       NAME/KEY:  modified_base
;       LOCATION:  3406..6470
;       OTHER INFORMATION:  /mod_base= OTHER
;       OTHER INFORMATION:  /note= "K = G or T"
;     FEATURE:
;       NAME/KEY:  modified_base
;       LOCATION:  3564..7896
;       OTHER INFORMATION:  /mod_base= OTHER
;       OTHER INFORMATION:  /note= "M = A or C"
;     FEATURE:
;       NAME/KEY:  modified_base
;       LOCATION:  3497..3607
;       OTHER INFORMATION:  /mod_base= OTHER
;       OTHER INFORMATION:  /note= "N = A or C or G or T"
;     FEATURE:
;       NAME/KEY:  modified_base
;       LOCATION:  3479..6422
;       OTHER INFORMATION:  /mod_base= OTHER
;       OTHER INFORMATION:  /note= "R = A or G"
;     FEATURE:
;       NAME/KEY:  modified_base
;       LOCATION:  3405..6871
;       OTHER INFORMATION:  /mod_base= OTHER
;       OTHER INFORMATION:  /note= "S = C or G"
      FEATURE:
        NAME/KEY:  modified_base
;       LOCATION:  3457..9998
;       OTHER INFORMATION:  /mod_base= OTHER
;       OTHER INFORMATION:  /note= "W = A or T"
;     FEATURE:
;       NAME/KEY:  modified_base
;       LOCATION:  3595..9999
;       OTHER INFORMATION:  /mod_base= OTHER
;       OTHER INFORMATION:  /note= "Y = C or T"
US-08-772-440-33


  Query Match             49.7%;  Score 14.4;  DB 3;  Length 10409;
  Best Local Similarity   47.6%;  Pred. No. 22;
  Matches   10;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       5 auncuuunnguaagcccnang 25
           |: |:::  |:| |||| | |
Db    8023 ATGCTTTGGGTATGCCCCAGG 8043

RESULT  11
US-08-008-216-19
; Sequence 19, Application US/08008216
```

```
; Patent No. 5366887
;  GENERAL INFORMATION:
;    APPLICANT:  Slightom, Jerry L.
;    APPLICANT:  Tepfer, David A.
;    TITLE OF INVENTION:  Ri T-DNA Promoters
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  GRAY, CARY, AMES & FRYE
;      STREET:  401 B Street, Suite 1700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-4297
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/008,216
;      FILING DATE:  25-JAN-1993
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 06/725,368
;      FILING DATE:  22-APR-1985
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Barnhorst, Marnie W.
;      REGISTRATION NUMBER:  36,740
;      REFERENCE/DOCKET NUMBER:  P1020US1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 699-2700
;      TELEFAX:  (619) 236-1048
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21126 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Agrobacterium rhizogenes
;      STRAIN:  STRAIN A4
;    IMMEDIATE SOURCE:
;      LIBRARY:  CONVOLVULUS ARVENSIS PLANT CELLS
;      CLONE:  CLONE 7
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  complement (937..2262)
;      OTHER INFORMATION:  /label= ORF1SUBSEQUENCE
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  complement (2649..3458)
;      OTHER INFORMATION:  /label= ORF2SUBSEQUENCE
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  3726..4799
;      OTHER INFORMATION:  /label= ORF3SUBSEQUENCE
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  complement (4041..4400)
;      OTHER INFORMATION:  /label= ORF4SUBSEQUENCE
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  complement (4607..4918)
;      OTHER INFORMATION:  /label= ORF5SUBSEQUENCE
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  5143..6216
;      OTHER INFORMATION:  /label= ORF6SUBSEQUENCE
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  complement (5071..5643)
```

```
;       OTHER INFORMATION:  /label= ORF7SUBSEQUENCE
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  6609..8888
;       OTHER INFORMATION:  /label= ORF8SUBSEQUENCE
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  complement (6576..6830)
;       OTHER INFORMATION:  /label= ORF9SUBSEQUENCE
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  9748..10044
;       OTHER INFORMATION:  /label= ORF10SUBSEQUENC
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  complement (10509..11282)
;       OTHER INFORMATION:  /label= ORF11SUBSEQUENC
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  12466..13002
;       OTHER INFORMATION:  /label= ORF12SUBSEQUENC
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  13723..14319
;       OTHER INFORMATION:  /label= ORF13SUBSEQUENC
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  15659..16210
;       OTHER INFORMATION:  /label= ORF14SUBSEQUENC
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  complement (16517..17545)
;       OTHER INFORMATION:  /label= ORF15SUBSEQUENC
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  complement (17737..18189)
;       OTHER INFORMATION:  /label= ORF16SUBSEQUENC
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  complement (18177..18743)
;       OTHER INFORMATION:  /label= ORF17SUBSEQUENC
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  complement (19031..19390)
;       OTHER INFORMATION:  /label= ORF18SUBSEQUENC
US-08-008-216-19


 Query Match             49.7%;  Score 14.4;  DB 1;  Length 21126;
 Best Local Similarity   47.6%;  Pred. No. 25;
 Matches   10;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       5 auncuuunnguaagcccnang 25
           |: |:::  |:||| || | |
Db   20133 ATTCTTTCGGTAAGACCGAAG 20153


RESULT  12
US-08-459-569-19
; Sequence 19, Application US/08459569
; Patent No. 5543501
;  GENERAL INFORMATION:
;    APPLICANT:  Slightom, Jerry L.
;    APPLICANT:  Tepfer, David A.
;    TITLE OF INVENTION:  Ri T-DNA Promoters
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  GRAY, CARY, AMES & FRYE
;      STREET:  401 B Street, Suite 1700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-4297
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/459,569
;      FILING DATE:  02-JUN-1995
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/008,216
;      FILING DATE:  25-JAN-1993
;      APPLICATION NUMBER:  US 06/725,368
;      FILING DATE:  22-APR-1985
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Barnhorst, Marnie W.
;      REGISTRATION NUMBER:  36,740
;      REFERENCE/DOCKET NUMBER:  P1020US1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 699-2700
;      TELEFAX:  (619) 236-1048
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21126 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Agrobacterium rhizogenes
;      STRAIN:  STRAIN A4
;    IMMEDIATE SOURCE:
;      LIBRARY:  CONVOLVULUS ARVENSIS PLANT CELLS
;      CLONE:  CLONE 7
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  complement (937..2262)
;      OTHER INFORMATION:  /label= ORF1SUBSEQUENCE
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  complement (2649..3458)
;      OTHER INFORMATION:  /label= ORF2SUBSEQUENCE
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  3726..4799
;      OTHER INFORMATION:  /label= ORF3SUBSEQUENCE
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  complement (4041..4400)
;      OTHER INFORMATION:  /label= ORF4SUBSEQUENCE
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  complement (4607..4918)
;      OTHER INFORMATION:  /label= ORF5SUBSEQUENCE
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  5143..6216
;      OTHER INFORMATION:  /label= ORF6SUBSEQUENCE
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  complement (5071..5643)
;      OTHER INFORMATION:  /label= ORF7SUBSEQUENCE
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  6609..8888
;      OTHER INFORMATION:  /label= ORF8SUBSEQUENCE
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  complement (6576..6830)
;      OTHER INFORMATION:  /label= ORF9SUBSEQUENCE
;    FEATURE:
;      NAME/KEY:  misc_feature
```

```
;       LOCATION:  9748..10044
;       OTHER INFORMATION:  /label= ORF10SUBSEQUENC
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  complement (10509..11282)
;       OTHER INFORMATION:  /label= ORF11SUBSEQUENC
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  12466..13002
;       OTHER INFORMATION:  /label= ORF12SUBSEQUENC
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  13723..14319
;       OTHER INFORMATION:  /label= ORF13SUBSEQUENC
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  15659..16210
;       OTHER INFORMATION:  /label= ORF14SUBSEQUENC
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  complement (16517..17545)
;       OTHER INFORMATION:  /label= ORF15SUBSEQUENC
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  complement (17737..18189)
;       OTHER INFORMATION:  /label= ORF16SUBSEQUENC
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  complement (18177..18743)
;       OTHER INFORMATION:  /label= ORF17SUBSEQUENC
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  complement (19031..19390)
;       OTHER INFORMATION:  /label= ORF18SUBSEQUENC
US-08-459-569-19


  Query Match             49.7%;  Score 14.4;  DB 1;  Length 21126;
  Best Local Similarity   47.6%;  Pred. No. 25;
  Matches   10;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       5 auncuuunnguaagcccnang 25
           |: |:::  |:||| || | |
Db   20133 ATTCTTTCGGTAAGACCGAAG 20153


RESULT  13
US-08-458-831-19
; Sequence 19, Application US/08458831
; Patent No. 5824866
;  GENERAL INFORMATION:
;    APPLICANT:  Slightom, Jerry L.
;    APPLICANT:  Tepfer, David A.
;    TITLE OF INVENTION:  Ri T-DNA Promoters
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  GRAY, CARY, AMES & FRYE
;      STREET:  401 B Street, Suite 1700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-4297
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/458,831
;      FILING DATE:  02-JUN-1995
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/008,216
;      FILING DATE:  25-JAN-1993
;      APPLICATION NUMBER:  US 06/725,368
;      FILING DATE:  22-APR-1985
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Barnhorst, Marnie W.
;      REGISTRATION NUMBER:  36,740
;      REFERENCE/DOCKET NUMBER:  P1020US1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 699-2700
;      TELEFAX:  (619) 236-1048
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21126 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Agrobacterium rhizogenes
;      STRAIN:  STRAIN A4
;    IMMEDIATE SOURCE:
;      LIBRARY:  CONVOLVULUS ARVENSIS PLANT CELLS
;      CLONE:  CLONE 7
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  complement (937..2262)
;      OTHER INFORMATION:  /label= ORF1SUBSEQUENCE
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  complement (2649..3458)
;      OTHER INFORMATION:  /label= ORF2SUBSEQUENCE
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  3726..4799
;      OTHER INFORMATION:  /label= ORF3SUBSEQUENCE
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  complement (4041..4400)
;      OTHER INFORMATION:  /label= ORF4SUBSEQUENCE
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  complement (4607..4918)
;      OTHER INFORMATION:  /label= ORF5SUBSEQUENCE
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  5143..6216
;      OTHER INFORMATION:  /label= ORF6SUBSEQUENCE
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  complement (5071..5643)
;      OTHER INFORMATION:  /label= ORF7SUBSEQUENCE
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  6609..8888
;      OTHER INFORMATION:  /label= ORF8SUBSEQUENCE
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  complement (6576..6830)
;      OTHER INFORMATION:  /label= ORF9SUBSEQUENCE
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  9748..10044
;      OTHER INFORMATION:  /label= ORF10SUBSEQUENC
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  complement (10509..11282)
;      OTHER INFORMATION:  /label= ORF11SUBSEQUENC
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  12466..13002
;      OTHER INFORMATION:  /label= ORF12SUBSEQUENC
;    FEATURE:
```

```
;       NAME/KEY:  misc_feature
;       LOCATION:  13723..14319
;       OTHER INFORMATION:  /label= ORF13SUBSEQUENC
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  15659..16210
;       OTHER INFORMATION:  /label= ORF14SUBSEQUENC
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  complement (16517..17545)
;       OTHER INFORMATION:  /label= ORF15SUBSEQUENC
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  complement (17737..18189)
;       OTHER INFORMATION:  /label= ORF16SUBSEQUENC
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  complement (18177..18743)
;       OTHER INFORMATION:  /label= ORF17SUBSEQUENC
      FEATURE:
        NAME/KEY:  misc_feature
        LOCATION:  complement (19031..19390)
;       OTHER INFORMATION:  /label= ORF18SUBSEQUENC
US-08-458-831-19


  Query Match             49.7%;  Score 14.4;  DB 1;  Length 21126;
  Best Local Similarity   47.6%;  Pred. No. 25;
  Matches   10;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       5 auncuuunnguaagcccnang 25
           |: |:::  |:||| || | |
Db  20133 ATTCTTTCGGTAAGACCGAAG 20153


RESULT  14
US-08-121-713D-61
; Sequence 61, Application US/08121713D
; Patent No. 5639856
;  GENERAL INFORMATION:
;    APPLICANT:  Goodman, Corey S.
;    APPLICANT:  Kolodkin, Alex L.
;    APPLICANT:  Matthes, David
;    APPLICANT:  Bentley, David R.
;    APPLICANT:  O'Connor, Timothy
;    TITLE OF INVENTION:  The Semaphorin Gene Family
;    NUMBER OF SEQUENCES:  100
     CORRESPONDENCE ADDRESS:
       ADDRESSEE:  SCIENCE & TECHNOLOGY LAW GROUP
       STREET:  268 Bush Street, Suite 3200
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/121,713D
;      FILING DATE:  13-SEP-1993
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Osman, Richard A.
;      REGISTRATION NUMBER:  36,627
;      REFERENCE/DOCKET NUMBER:  B94-002-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415)343-4341
;      TELEFAX:  (415) 343-4342
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  61:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2670 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  268..2439
US-08-121-713D-61


  Query Match             47.6%;  Score 13.8;  DB 1;  Length 2670;
  Best Local Similarity   50.0%;  Pred. No. 37;
  Matches   11;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       4 gauncuuunnguaagcccnang 25
           ||   :::  |:|||||| | |
Db  2416 GAGATTTTCCGTAAGCCCAATG 2437


RESULT  15
US-08-835-268-61
; Sequence 61, Application US/08835268
; Patent No. 5807826
;  GENERAL INFORMATION:
;    APPLICANT:  Goodman, Corey S.
;    APPLICANT:  Kolodkin, Alex L.
;    APPLICANT:  Matthes, David
;    APPLICANT:  Bentley, David R.
;    APPLICANT:  O'Connor, Timothy
;    TITLE OF INVENTION:  The Semaphorin Gene Family
;    NUMBER OF SEQUENCES:  100
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SCIENCE & TECHNOLOGY LAW GROUP
;      STREET:  268 Bush Street, Suite 3200
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/835,268
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/121,713
;      FILING DATE:  13-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Osman, Richard A.
;      REGISTRATION NUMBER:  36,627
;      REFERENCE/DOCKET NUMBER:  B94-002-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415)343-4341
;      TELEFAX:  (415) 343-4342
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  61:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2670 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  268..2439
US-08-835-268-61


  Query Match             47.6%;  Score 13.8;  DB 1;  Length 2670;
```

Best Local Similarity 50.0%; Pred. No. 37;
Matches 11; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

```
Qy       4 gauncuuunnguaagcccnang 25
           ||   :::  |:|||||| | |
Db    2416 GAGATTTTCCGTAAGCCCAATG 2437
```

Search completed: October 2, 2001, 05:01:46
Job time: 5313 sec

GenCore version 4.5


OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 03:31:53 ; Search time 1315.38 Seconds
(without alignments)
341.015 Million cell updates/sec

Title: US-09-310-844B-23
Perfect score: 29
Sequence: 1 nnngauncuuunnguaagcccnangngnn 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: em_ba1:*
17: em_ba2:*
18: em_fun:*
19: em_htgo_hum:*
20: em_htgo_inv:*
21: em_htgo_rod:*
22: em_htg_hum1:*
23: em_htg_hum2:*
24: em_htg_hum3:*
25: em_htg_hum4:*
26: em_htg_hum5:*
27: em_htg_hum6:*
28: em_htg_hum7:*
29: em_htg_hum8:*
30: em_htg_inv1:*
31: em_htg_inv2:*
32: em_htg_other:*
33: em_htg_rod:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_om:*
43: em_or:*
44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pl:*
48: em_ro:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_vi:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_vi1:*
59: gb_vi2:*
60: gb_htg1:*
61: gb_htg2:*
62: gb_htg3:*
63: gb_htg4:*
64: gb_htg5:*
65: gb_htg6:*
66: gb_htg7:*
67: gb_htg8:*
68: gb_htg9:*
69: gb_htg10:*
70: gb_htg11:*
71: gb_htg12:*
72: gb_htg13:*
73: gb_htg14:*
74: gb_htg15:*
75: gb_htg16:*
76: gb_htg17:*
77: gb_htg18:*
78: gb_htg19:*
79: gb_htg20:*
80: gb_htg21:*
81: gb_htg22:*
82: gb_htg23:*
83: gb_htg24:*
84: gb_htg25:*
85: gb_pr1:*
86: gb_pr2:*
87: gb_pr3:*
88: gb_pr4:*
89: gb_pr5:*
90: gb_pr6:*
91: gb_pr7:*
92: gb_pr8:*
93: gb_pr9:*
94: gb_ro1:*
95: gb_ro2:*
96: gb_in4:*
97: gb_pr10:*
98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 18 | 62.1 | 209 | 10 | I82323 | I82323 Sequence 8 |
| 2 | 18 | 62.1 | 292 | 54 | G06364 | G06364 human STS W |
| 3 | 18 | 62.1 | 664 | 7 | MAU79187 | U79187 Mirounga an |
| 4 | 18 | 62.1 | 689 | 9 | A17717 | A17717 Plasmid pTG |
| 5 | 18 | 62.1 | 698 | 9 | A02159 | A02159 Synthetic g |
| 6 | 18 | 62.1 | 698 | 9 | A04681 | A04681 Synthetic g |
| 7 | 18 | 62.1 | 698 | 9 | A14089 | A14089 Synthetic D |
| 8 | 18 | 62.1 | 722 | 10 | I03416 | I03416 Sequence 2 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 9 | 18 | 62.1 | 730 | 91 | GIBIL2 | K03174 Ape (gibbon |
| 10 | 18 | 62.1 | 740 | 95 | RATIL2 | M22899 Rat interle |
| 11 | 18 | 62.1 | 750 | 10 | I01391 | I01391 Sequence 3 |
| 12 | 18 | 62.1 | 764 | 7 | BOVIL2 | M12791 Bovine inte |
| 13 | 18 | 62.1 | 769 | 10 | E01460 | E01460 cDNA encodi |
| 14 | 18 | 62.1 | 769 | 10 | E02201 | E02201 DNA encodin |
| 15 | 18 | 62.1 | 773 | 7 | CFU28141 | U28141 Canis famil |
| 16 | 18 | 62.1 | 779 | 7 | CATIL2A | L19402 Felis catus |
| 17 | 18 | 62.1 | 782 | 10 | I01101 | I01101 Sequence 2 |
| 18 | 18 | 62.1 | 784 | 9 | A06879 | A06879 Artificial |
| 19 | 18 | 62.1 | 784 | 9 | A14095 | A14095 Synthetic D |
| 20 | 18 | 62.1 | 784 | 9 | A17732 | A17732 Plasmid pTG |
| 21 | 18 | 62.1 | 784 | 10 | E00250 | E00250 DNA coding |
| 22 | 18 | 62.1 | 784 | 10 | I08366 | I08366 Sequence 2 |
| 23 | 18 | 62.1 | 788 | 10 | I07944 | I07944 Sequence 1 |
| 24 | 18 | 62.1 | 790 | 10 | I04498 | I04498 Sequence 1 |
| 25 | 18 | 62.1 | 790 | 10 | I08377 | I08377 Sequence 1 |
| 26 | 18 | 62.1 | 791 | 8 | MGA7463 | AJ007463 Meleagris |
| 27 | 18 | 62.1 | 794 | 10 | E00210 | E00210 cDNA encodi |
| 28 | 18 | 62.1 | 794 | 10 | E00211 | E00211 cDNA encodi |
| 29 | 18 | 62.1 | 794 | 10 | E00215 | E00215 cDNA encodi |
| 30 | 18 | 62.1 | 794 | 10 | E02011 | E02011 cDNA encodi |
| 31 | 18 | 62.1 | 794 | 10 | E02018 | E02018 cDNA encodi |
| 32 | 18 | 62.1 | 795 | 10 | E00214 | E00214 cDNA encodi |
| 33 | 18 | 62.1 | 795 | 10 | E00267 | E00267 DNA sequenc |
| 34 | 18 | 62.1 | 795 | 10 | E00271 | E00271 DNA sequenc |
| 35 | 18 | 62.1 | 798 | 8 | AF033563 | AF033563 Gallus ga |
| 36 | 18 | 62.1 | 800 | 8 | AF017645 | AF017645 Gallus ga |
| 37 | 18 | 62.1 | 800 | 10 | E59473 | E59473 Novel cytok |
| 38 | 18 | 62.1 | 801 | 9 | A04961 | A04961 Artificial |
| 39 | 18 | 62.1 | 801 | 9 | A06759 | A06759 Artificial |
| 40 | 18 | 62.1 | 801 | 10 | E00336 | E00336 Human inter |
| 41 | 18 | 62.1 | 801 | 10 | I01197 | I01197 Sequence 1 |
| 42 | 18 | 62.1 | 801 | 10 | I04318 | I04318 Sequence 3 |
| 43 | 18 | 62.1 | 801 | 93 | HSIL02 | V00564 Human mRNA |
| 44 | 18 | 62.1 | 810 | 10 | E00978 | E00978 cDNA sequen |
| 45 | 18 | 62.1 | 810 | 10 | I01387 | I01387 Sequence 1 |

ALIGNMENTS

```
RESULT   1
I82323
LOCUS       I82323         209 bp    DNA             PAT       10-JUN-1998
DEFINITION  Sequence 8 from patent US 5712126.
ACCESSION   I82323
VERSION     I82323.1  GI:3210620
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 209)
  AUTHORS   Weissman,S.M. and Prashar,Y.
  TITLE     Analysis of gene expression by display of 3-end restriction
            fragments of CDNA
  JOURNAL   Patent: US 5712126-A 8 27-JAN-1998;
FEATURES             Location/Qualifiers
     source          1. .209
                     /organism="unknown"
BASE COUNT       56 a     30 c     33 g     90 t
ORIGIN


  Query Match             62.1%;  Score 18;  DB 10;  Length 209;
  Best Local Similarity   54.2%;  Pred. No. 1.5;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      4 gauncuuunnguaagcccnangng 27
          ||: |::: |:|||||| | | |
Db     36 GATTCTTTTTGTAAGCCCTAGGGG 59
```

```
RESULT   2
G06364
LOCUS       G06364         292 bp    DNA             STS       19-OCT-1995
DEFINITION  human STS WI-7035.
ACCESSION   G06364
VERSION     G06364.1  GI:859609
KEYWORDS    STS sequence; primer; sequence tagged site.
SOURCE      human STSs derived from sequences in dbEST and the Unigene
            collection.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata;
            Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates;
            Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 292)
  AUTHORS   Hudson,T.
  TITLE     Whitehead Institute/MIT Center for Genome Research; Physically
            Mapped ESTs
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Thomas Hudson
            Whitehead Institute/MIT Center for Genome Research
            Whitehead Institute for Biomedical Research
            9 Cambridge Center, Cambridge MA 02142 USA
            Tel: 617 252 1900
            Fax: 617 252 1902
            Email: thudson@genome.wi.mit.edu

            Primer A: TAATTAAGTGCTTCCCACTTAAAAC
            Primer B: ATTTTGGGATAAATAAGTGAAACCA
            STS size: 200
            PCR Profile:
                  Presoak:
                  Denaturation:
                  Annealing: 56 degrees C
                  Polymerization:
                  PCR Cycles: 35
                  Thermal Cycler:
            Protocol:
                Template: 10 ng
                Primer: each 5 pM
                dNTPs: each 4 nM
                Taq Polymerase: 0.025 units/ul
                Total Vol: 20 ul

            Buffer:
                MgCl2: 1.5 mM
                KCl: 50 mM
                Tris-HCL: 10 mM
                pH: 9.3

            Prepared with primer pairs derived from V00564 -- Unigene.
FEATURES             Location/Qualifiers
     source          1. .292
                     /organism="Homo sapiens"
     STS             1. .200
     primer_bind     1. .25
     primer_bind     complement(176. .200)
BASE COUNT       92 a     32 c     32 g    124 t     12 others
ORIGIN


  Query Match             62.1%;  Score 18;  DB 54;  Length 292;
  Best Local Similarity   54.2%;  Pred. No. 1.6;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      4 gauncuuunnguaagcccnangng 27
          ||: |::: |:|||||| | | |
Db    144 GATTCTTTTTGTAAGCCCTAGGGG 167


RESULT   3
MAU79187
```

```
LOCUS       MAU79187      664 bp    mRNA            MAM       03-APR-1998
DEFINITION  Mirounga angustirostris interleukin 2 precursor mRNA, complete cds.
ACCESSION   U79187
VERSION     U79187.1  GI:3015531
KEYWORDS    .
SOURCE      northern elephant seal.
  ORGANISM  Mirounga angustirostris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Pinnipedia; Phocidae; Mirounga.
REFERENCE   1  (bases 1 to 664)
  AUTHORS   Shoda,L.K.M., Brown,W.C. and Rice-Ficht,A.C.
  TITLE     Sequence and characterization of phocine interleukin 2
  JOURNAL   J. Wildl. Dis. 34 (1), 81-90 (1998)
  MEDLINE   98136706
REFERENCE   2  (bases 1 to 664)
  AUTHORS   Shoda,L.K.M., Brown,W.C. and Rice-Ficht,A.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-NOV-1996) Veterinary Microbiology & Pathology,
            Washington State University, P.O. Box 647040, Pullman, WA
            99164-7040, USA
  ATURES              Location/Qualifiers
     source           1. .664
                      /organism="Mirounga angustirostris"
                      /db_xref="taxon:9716"
     sig_peptide      24. .83
     CDS              24. .488
                      /codon_start=1
                      /product="interleukin 2 precursor"
                      /protein_id="AAC12258.1"
                      /db_xref="GI:3015532"
                      /translation="MCKMQLLSCIALSLVLVANSAPTTSSTKETQQQLEQLLLDLRLL
                      LNGVNNYEDPKLSRMLTFKFYTPKKATELTHLQCLAEELKPLEEVLYLAQSKNFHLTD
                      IKELMSNINVTLLKLKGSETRFKCEYDDETATITEFLNKWITFCQSIFSTLT"
     mat_peptide      84. .485
                      /product="interleukin 2"
BASE COUNT      219 a    131 c    106 g    208 t
ORIGIN


  Query Match             62.1%;  Score 18;  DB 7;  Length 664;
  Best Local Similarity   54.2%;  Pred. No. 1.9;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      4 gauncuuunnguaagcccnangng 27
          ||: |::: |:|||||| | | |
Db    635 GATTCTTTTTGTAAGCCCTAGGCG 658


:SULT   4
17717
LOCUS       A17717        689 bp    mRNA            PAT       17-MAY-1994
DEFINITION  Plasmid pTG 26 mRNA for human IL-2.
ACCESSION   A17717
VERSION     A17717.1  GI:513949
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  Plasmid unidentified
            unclassified.
REFERENCE   1  (bases 1 to 689)
  AUTHORS   Kieny,M.P., Sondermeyer,P. and Lecocq,J.P.
  TITLE     Expression of human IL-2 in mammalian cells by a recombined pox-
            virus
  JOURNAL   Patent: EP 0206939-A 1 30-DEC-1986;
            TRANSGENE S.A
FEATURES              Location/Qualifiers
     source           1. .689
                      /organism="unidentified"
                      /plasmid="pTG26"
                      /db_xref="taxon:32644"
     CDS              1. .384
                      /partial
                      /codon_start=1
                      /product="human interleukin 2"
                      /protein_id="CAA01347.1"
                      /db_xref="GI:513950"
                      /translation="TKKTQLQLEHLLLDLQMILNGINNYKNPKLTRMLTFKFYMPKKA
                      TELKHLQCLEEELKPLEEVLNLAQSKNFHLRPRDLISNINVIVLELKGSETTFMCEYA
                      DETATIVEFLNRWITFCQSIISTLT"
BASE COUNT      251 a    117 c     99 g    222 t
ORIGIN


  Query Match             62.1%;  Score 18;  DB 9;  Length 689;
  Best Local Similarity   54.2%;  Pred. No. 1.9;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      4 gauncuuunnguaagcccnangng 27
          ||: |::: |:|||||| | | |
Db    528 GATTCTTTTTGTAAGCCCTAGGGG 551


RESULT   5
A02159
LOCUS       A02159        698 bp    DNA             PAT       21-MAY-1993
DEFINITION  Synthetic gene for interleukin 2 (IL-2) (partial).
ACCESSION   A02159
VERSION     A02159.1  GI:412310
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 698)
  AUTHORS   .
  TITLE     VECTOR FOR THE EXPRESSION IN YEASTS OF INTERLEUKINE-2, TRANSFORMED
            YEASTS AND METHOD FOR PREPARING INTERLEUKINE-2
  JOURNAL   Patent: WO 8503723-A 6 29-AUG-1985;
FEATURES              Location/Qualifiers
     source           1. .698
                      /organism="synthetic construct"
                      /db_xref="taxon:32630"
     CDS              10. .393
                      /partial
                      /codon_start=1
                      /transl_table=11
                      /product="interleukin 2 (IL-2)"
                      /protein_id="CAA00227.1"
                      /db_xref="GI:412311"
                      /translation="TKKTQLQLEHLLLDLQMILNGINNYKNPKLTRMLTFKFYMPKKA
                      TELKHLQCLEEELKPLEEVLNLAQSKNFHLRPRDLISNINVIVLELKGSETTFMCEYA
                      DETATIVEFLNRWITFCQSIISTLT"
BASE COUNT      251 a    117 c    107 g    222 t      1 others
ORIGIN


  Query Match             62.1%;  Score 18;  DB 9;  Length 698;
  Best Local Similarity   54.2%;  Pred. No. 1.9;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      4 gauncuuunnguaagcccnangng 27
          ||: |::: |:|||||| | | |
Db    537 GATTCTTTTTGTAAGCCCTAGGGG 560


RESULT   6
A04681
LOCUS       A04681        698 bp    DNA             PAT       24-MAY-1993
DEFINITION  Synthetic gene for interleukin 2 (IL-2) (partial).
ACCESSION   A04681
VERSION     A04681.1  GI:412434
KEYWORDS    interleukin 2.
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 698)
  AUTHORS   Lemoine,Y., Sondermeyer,P., Loison,G., Aigle,M. and Lecocq,J.P.
  TITLE     Yeast-expression vectors for interleukin-2, transformed yeasts and
```

```
          process for the preparation of interleukin-2
  JOURNAL   Patent: EP 0152358-A 5 21-AUG-1985;
          TRANSGENE S.A
FEATURES             Location/Qualifiers
     source          1. .698
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     CDS             10. .393
                     /partial
                     /codon_start=1
                     /transl_table=11
                     /product="interleukin 2 (IL-2)"
                     /protein_id="CAA00377.1"
                     /db_xref="GI:412435"
                     /translation="TKKTQLQLEHLLLDLQMILNAINNYKNPKLTRMLTFKFYMPKKA
                     TELKHLQCLEEELKPLEEVLNLAQSKNFHLRPRDLISNINVIVLELKGSETTFMCEYA
                     DETATIVEFLNRWITFCQSIISTLT"
BASE COUNT      251 a    118 c    106 g    222 t      1 others
ORIGIN


  Query Match             62.1%;  Score 18;  DB 9;  Length 698;
  Best Local Similarity   54.2%;  Pred. No. 1.9;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      4 gauncuuunnguaagcccnangng 27
          ||: |::: |:|||||| | | |
Db    537 GATTCTTTTTGTAAGCCCTAGGGG 560


RESULT   7
A14089
LOCUS       A14089        698 bp    DNA             PAT       28-FEB-1994
DEFINITION  Synthetic DNA (pTG26) for human interleukin-2 (partial).
ACCESSION   A14089
VERSION     A14089.1  GI:491760
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
          artificial sequence.
REFERENCE   1  (bases 1 to 698)
  AUTHORS   .
  TITLE     EXPRESSION OF HUMAN IL-2 IN MAMMAL CELLS BY MEANS OF A RECOMBINED
          POXVIRUS
  JOURNAL   Patent: WO 8607610-A 1 31-DEC-1986;
FEATURES             Location/Qualifiers
     source          1. .698
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     CDS             10. .393
                     /partial
                     /codon_start=1
                     /transl_table=11
                     /product="interleukin 2"
                     /protein_id="CAA01145.1"
                     /db_xref="GI:491761"
                     /translation="TKKTQLQLEHLLLDLQMILNGINNYKNPKLTRMLTFKFYMPKKA
                     TELKHLQCLEEELKPLEEVLNLAQSKNFHLRPRDLISNINVIVLELKGSETTFMCEYA
                     DETATIVEFLNRWITFCQSIISTLT"
BASE COUNT      251 a    117 c    108 g    222 t
ORIGIN


  Query Match             62.1%;  Score 18;  DB 9;  Length 698;
  Best Local Similarity   54.2%;  Pred. No. 1.9;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      4 gauncuuunnguaagcccnangng 27
          ||: |::: |:|||||| | | |
Db    537 GATTCTTTTTGTAAGCCCTAGGGG 560


RESULT   8
I03416
LOCUS       I03416        722 bp ss-DNA             PAT       21-MAY-1993
DEFINITION  Sequence 2 from Patent US 4882282.
ACCESSION   I03416
VERSION     I03416.1  GI:270625
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
          Unclassified.
REFERENCE   1  (bases 1 to 722)
  AUTHORS   Anderson,D.M., Baker,P.E., Cantrell,M.A., Cerretti,D.P.,
          Cosman,D.J., Gimpel,S.D., Grabstein,K.H., Larsen,A.D. and
          McKereghan,K.N.
  TITLE     DNA sequences encoding bovine interleukin-2
  JOURNAL   Patent: US 4882282-A 2 21-NOV-1989;
          Immunex Corporation;
          Seattle, WA
FEATURES             Location/Qualifiers
     source          1. .722
                     /organism="unknown"
BASE COUNT      245 a    130 c    112 g    235 t
ORIGIN


  Query Match             62.1%;  Score 18;  DB 10;  Length 722;
  Best Local Similarity   54.2%;  Pred. No. 1.9;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      4 gauncuuunnguaagcccnangng 27
          ||: |::: |:|||||| | | |
Db    613 GATTCTTTTTGTAAGCCCTAGGGG 636


RESULT   9
GIBIL2
LOCUS       GIBIL2        730 bp    mRNA            PRI       27-APR-1993
DEFINITION  Ape (gibbon) interleukin 2 mRNA.
ACCESSION   K03174
VERSION     K03174.1  GI:177012
KEYWORDS    interleukin; interleukin 2.
SOURCE      Gibbon ('normal gibbon'), cDNA to mRNA, clone CM11.
  ORGANISM  Hylobates lar
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
REFERENCE   1  (bases 1 to 730)
  AUTHORS   Chen,S.J., Holbrook,N.J., Mitchell,K.F., Vallone,C.A.,
          Greengard,J.S., Crabtree,G.R. and Lin,Y.
  TITLE     A viral long terminal repeat in the interleukin 2 gene of a cell
          line that constitutively produces interleukin 2
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 82, 7284-7288 (1985)
  MEDLINE   86042650
COMMENT     [1] also sequenced the IL-2 mRNA from gibbon cell line MLA 144 (see
          separate entry).
FEATURES             Location/Qualifiers
     source          1. .730
                     /organism="Hylobates lar"
                     /db_xref="taxon:9580"
     mRNA            <1. .>729
                     /note="IL-2 mRNA"
     CDS             25. .486
                     /note="interleukin-2"
                     /codon_start=1
                     /protein_id="AAA35453.1"
                     /db_xref="GI:177013"
                     /translation="MYRMQLLSCIALSLALVTNSAPTSSSTKKTQLQLEHLLLDLQMI
                     LNGINNYKNPKLTRMLTFKFYMPKKATELKHLQCLEEELKPLEEVLNLAQSKNFHLRP
                     RDLISNINVIVLELKGSETTFMCEYADETATIVEFLNRWITFCQSIISTLT"
     variation       241
                     /note="c in CM11; t in CM9"
BASE COUNT      247 a    138 c    110 g    235 t
ORIGIN      180 bp upstream of HinfI site.
```

```
  Query Match            62.1%;  Score 18;  DB 91;  Length 730;
  Best Local Similarity  54.2%;  Pred. No. 1.9;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      4 gauncuuunnguaagcccnangng 27
          ||: |::: |:|||||| | | |
Db    630 GATTCTTTTTGTAAGCCCTAGGGG 653


RESULT  10
RATIL2
LOCUS       RATIL2        740 bp    mRNA            ROD       27-APR-1993
DEFINITION  Rat interleukin 2 mRNA, complete cds.
ACCESSION   M22899
VERSION     M22899.1  GI:204909
KEYWORDS    interleukin 2.
SOURCE      Rat activated T-lymphocytes, cDNA to mRNA, clone pRIL-2.8.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 740)
  AUTHORS   McKnight,A.J., Mason,D.W. and Barclay,A.N.
  TITLE     Sequence of rat interleukin 2 and anomalous binding of a mouse
            interleukin 2 cDNA probe to rat MHC class II-associated invariant
            chain mRNA
  JOURNAL   Immunogenetics 30, 145-147 (1989)
  MEDLINE   89339608
COMMENT     Draft entry and computer-readable sequence for [1] kindly submitted
            by A.J.McKnight, 08-MAR-1989.
FEATURES             Location/Qualifiers
     source          1. .740
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
     CDS             15. .482
                     /note="interleukin 2 precursor"
                     /codon_start=1
                     /protein_id="AAA41427.1"
                     /db_xref="GI:204910"
                     /translation="MYSMQLASCVALTLVLLVNSAPTSSPAKETQQHLEQLLLDLQVL
                     LRGIDNYKNLKLPMMLTFKFYLPKQATELKHLQCLENELGALQRVLDLTQSKSFHLED
                     AGNFISNIRVTVVKLKGSENKFECQFDDEPATVVEFLRRWIAICQSIISTMTQ"
     sig_peptide     15. .74
                     /note="interleukin 2 signal peptide"
     mat_peptide     75. .479
                     /note="interleukin 2"
     polyA_signal    713. .718
                     /note="putative"
BASE COUNT      228 a    141 c    143 g    228 t
ORIGIN


  Query Match            62.1%;  Score 18;  DB 95;  Length 740;
  Best Local Similarity  54.2%;  Pred. No. 1.9;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      4 gauncuuunnguaagcccnangng 27
          ||: |::: |:|||||| | | |
Db    615 GATTCTTTTTGTAAGCCCCAAGGG 638


RESULT  11
I01391
LOCUS       I01391        750 bp ss-DNA             PAT       21-MAY-1993
DEFINITION  Sequence 3 from Patent US 4851512.
ACCESSION   I01391
VERSION     I01391.1  GI:270228
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 750)
  AUTHORS   Miyaji,H. and Itoh,S.
  TITLE     Novel human interleukin-2 polypeptide derivative
  JOURNAL   Patent: US 4851512-A 3 25-JUL-1989;
            Kyowa Hakko Kogyo Co., Ltd.;
            Tokyo, ;
            JP;
FEATURES             Location/Qualifiers
     source          1. .750
                     /organism="unknown"
BASE COUNT      276 a    131 c    104 g    239 t
ORIGIN


  Query Match            62.1%;  Score 18;  DB 10;  Length 750;
  Best Local Similarity  54.2%;  Pred. No. 1.9;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      4 gauncuuunnguaagcccnangng 27
          ||: |::: |:|||||| | | |
Db    591 GATTCTTTTTGTAAGCCCTAGGGG 614


RESULT  12
BOVIL2
LOCUS       BOVIL2        764 bp    mRNA            MAM       27-APR-1993
DEFINITION  Bovine interleukin 2 (IL-2) mRNA, complete cds.
ACCESSION   M12791
VERSION     M12791.1  GI:163204
KEYWORDS    interleukin 2.
SOURCE      Bovine lymph node, cDNA to mRNA, clone pBIL2-4.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 764)
  AUTHORS   Cerretti,D.P., McKereghan,K., Larsen,A., Cantrell,M.A.,
            Anderson,D., Gillis,S., Cosman,D. and Baker,P.E.
  TITLE     Cloning, sequence, and expression of bovine interleukin 2
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 83, 3223-3227 (1986)
  MEDLINE   86205869
COMMENT     Draft entry and clean copy sequence for [1] kindly provided by
            D.Cerretti, 12-AUG-1986.
            There is probably only one copy of the interleukin 2 gene in the
            bovine genome.
FEATURES             Location/Qualifiers
     source          1. .764
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /tissue_type="lymph"
     mRNA            <1. .764
                     /gene="IL-2"
     gene            1. .764
                     /gene="IL-2"
     sig_peptide     18. .77
                     /gene="IL-2"
     CDS             18. .485
                     /gene="IL-2"
                     /note="prepeptide"
                     /codon_start=1
                     /product="interleukin 2"
                     /protein_id="AAA30586.1"
                     /db_xref="GI:163205"
                     /translation="MYKIQLLSCIALTLALVANGAPTSSSTGNTMKEVKSLLLDLQLL
                     LEKVKNPENLKLSRMHTFDFYVPKVNATELKHLKCLLEELKLLEEVLNLAPSKNLNPR
                     EIKDSMDNIKRIVLELQGSETRFTCEYDDATVNAVEFLNKWITFCQSIYSTMT"
     mat_peptide     78. .482
                     /gene="IL-2"
                     /product="interleukin 2"
BASE COUNT      257 a    133 c    123 g    251 t
ORIGIN      80 bp upstream of HgiAI site.


  Query Match            62.1%;  Score 18;  DB 7;  Length 764;
  Best Local Similarity  54.2%;  Pred. No. 1.9;
```

```
 Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      4 gauncuuunnguaagcccnangng 27
          ||: |::: |:|||||| | | |
Db    622 GATTCTTTTTGTAAGCCCTACGGG 645


RESULT  13
E01460
LOCUS       E01460        769 bp    RNA             PAT       29-SEP-1997
DEFINITION  cDNA encoding human interleukin 2.
ACCESSION   E01460
VERSION     E01460.1  GI:2169716
KEYWORDS    JP 1987265992-A/1.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 769)
  AUTHORS   Onomichi,K., Eto,Y. and Shibai,H. .
  TITLE     PRODUCTION OF VALUABLE SUBSTANCE USING HUMAN CELL
  JOURNAL   Patent: JP 1987265992-A 1 18-NOV-1987;
            AJINOMOTO CO INC
COMMENT     OS   Human
            PN   JP 1987265992-A/1
            PD   18-NOV-1987
            PF   12-MAY-1986 JP 1986107981
            PI   ONOMICHI KAZUYA, ETO YUZURU, SHIBAI HIROSHIRO PC
            C12P21/00,C12N15/00,C12P21/02,(C12P21/02,C12R1:91); CC
            strandedness: Double;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No;
            CC   *source: clone=pSDI;
            FH   Key             Location/Qualifiers
            FH
            FT   5'UTR           1. .2
            FT   CDS             3. .507
            FT                   /product='human interleukin 2' FT   3'UTR
                 508. .814
            FT   mat_peptide     3. .504
            FT                   /product='human interleukin 2'.
FEATURES              Location/Qualifiers
     source           1. .769
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
BASE COUNT       280 a    132 c    112 g    245 t
ORIGIN


  Query Match             62.1%;  Score 18;  DB 10;  Length 769;
  Best Local Similarity   54.2%;  Pred. No. 1.9;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      4 gauncuuunnguaagcccnangng 27
          ||: |::: |:|||||| | | |
Db    608 GATTCTTTTTGTAAGCCCTAGGGG 631


RESULT  14
E02201
LOCUS       E02201        769 bp    RNA             PAT       29-SEP-1997
DEFINITION  DNA encoding interleukin2 (IL-2).
ACCESSION   E02201
VERSION     E02201.1  GI:2170439
KEYWORDS    JP 1990009388-A/2.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 769)
  AUTHORS   Murata,M., Eto,Y. and Shibai,H. .
  TITLE     PRODUCTION OF PHYSIOLOGICALLY ACTIVE PROTEIN
  JOURNAL   Patent: JP 1990009388-A 2 12-JAN-1990;
            AJINOMOTO CO INC
COMMENT     OS   Homo sapiens
            PN   JP 1990009388-A/2
            PD   12-JAN-1990
            PF   08-JUL-1988 JP 1988170142
            PR   09-MAR-1988 JP 88P    55270
            PI   MURATA MASAHIRO, ETO YUZURU, SHIBAI HIROSHIRO PC
            C12P21/00,C12N15/12,C12N15/26,C12P21/02,(C12P21/02,C12R1:91); CC
            strandedness: Double;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No;
            CC   *source: library=pSD(I)/IL-2;
            CC   Feature is identified by experimental;
            FH   Key             Location/Qualifiers
            FH
            FT   5'UTR           1. .2
            FT   CDS             3. .464
            FT                   /product='interleukin2'
            FT                   /note='IL-2'
            FT   mat_peptide     3. .461
            FT                   /product='interleukin2'
            FT                   /note='IL-2'
            FT   3'UTR           465. .769.
FEATURES              Location/Qualifiers
     source           1. .769
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
BASE COUNT       280 a    132 c    112 g    245 t
ORIGIN


  Query Match             62.1%;  Score 18;  DB 10;  Length 769;
  Best Local Similarity   54.2%;  Pred. No. 1.9;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      4 gauncuuunnguaagcccnangng 27
          ||: |::: |:|||||| | | |
Db    608 GATTCTTTTTGTAAGCCCTAGGGG 631


RESULT  15
CFU28141
LOCUS       CFU28141      773 bp    mRNA            MAM       29-JUN-1995
DEFINITION  Canis familiaris interleukin-2 mRNA, complete cds.
ACCESSION   U28141
VERSION     U28141.1  GI:881935
KEYWORDS    .
SOURCE      dog.
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1  (bases 1 to 773)
  AUTHORS   Somberg,R.L., Tipold,A., Henthorn,P.S. and Felsburg,P.J.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 773)
  AUTHORS   Somberg,R.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-MAY-1995) Richard L. Somberg, Clinical Studies,
            University of Pennsylvania, 3850 Spruce Street, Philadelphia, PA
            19104, USA
FEATURES              Location/Qualifiers
     source           1. .773
                      /organism="Canis familiaris"
                      /db_xref="taxon:9615"
     CDS              39. .506
                      /codon_start=1
                      /product="interleukin-2"
                      /protein_id="AAA68969.1"
                      /db_xref="GI:881936"
                      /translation="MYKMQLLSCIALTLVLVANSAPITSSSTKETEQQMEQLLLDLQL
                      LLNGVNNYENPQLSRMLTFKFYTPKKATEFTHLQCLAEELKNLEEVLGLPQSKNVHLT
```

```
                     DTKELISNMNVTLLKLKGSETSYNCEYDDETATITEFLNKWITFCQSIFSTLT"
     polyA_site      773
                     /note="38 A nucleotides"
BASE COUNT      265 a    145 c    118 g    245 t
ORIGIN


  Query Match             62.1%;  Score 18;  DB 7;  Length 773;
  Best Local Similarity   54.2%;  Pred. No. 1.9;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       4 gauncuuunnguaagcccnangng 27
           ||: |::: |:|||||| | | |
Db     643 GATTCTTTTTGTAAGCCCTAGGGG 666


Search completed: October  2, 2001, 05:00:40
Job time: 5327 sec
```

GenCore version 4.5


OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 04:00:28 ; Search time 122.48 Seconds
(without alignments)
148.670 Million cell updates/sec

Title: US-09-310-844B-23
Perfect score: 29
Sequence: 1 nnngauncuuunnguaagcccnangngnn 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_0601:*
1: /SIDS1/gcgdata/geneseq/geneseqn/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 18 | 62.1 | 29 | 21 | AAA70827 | Molecular interact |
| 2 | 18 | 62.1 | 29 | 21 | AAA70828 | Molecular interact |
| 3 | 18 | 62.1 | 29 | 21 | AAA70829 | Molecular interact |
| 4 | 18 | 62.1 | 29 | 21 | AAA70830 | Molecular interact |
| 5 | 18 | 62.1 | 42 | 21 | AAA71113 | Molecular interact |
| 6 | 18 | 62.1 | 42 | 21 | AAA71114 | Molecular interact |
| 7 | 18 | 62.1 | 42 | 21 | AAA71115 | Molecular interact |
| 8 | 18 | 62.1 | 42 | 21 | AAA71116 | Molecular interact |
| 9 | 18 | 62.1 | 42 | 21 | AAA71118 | Molecular interact |
| 10 | 18 | 62.1 | 42 | 21 | AAA71119 | Molecular interact |
| 11 | 18 | 62.1 | 42 | 21 | AAA71120 | Molecular interact |
| 12 | 18 | 62.1 | 42 | 21 | AAA71121 | Molecular interact |
| 13 | 18 | 62.1 | 42 | 21 | AAA71123 | Molecular interact |
| 14 | 18 | 62.1 | 42 | 21 | AAA71124 | Molecular interact |
| 15 | 18 | 62.1 | 42 | 21 | AAA71126 | Molecular interact |
| 16 | 18 | 62.1 | 42 | 21 | AAA71127 | Molecular interact |
| 17 | 18 | 62.1 | 42 | 21 | AAA71128 | Molecular interact |
| 18 | 18 | 62.1 | 42 | 21 | AAA71129 | Molecular interact |
| 19 | 18 | 62.1 | 42 | 21 | AAA71131 | Molecular interact |
| 20 | 18 | 62.1 | 42 | 21 | AAA71132 | Molecular interact |
| 21 | 18 | 62.1 | 44 | 21 | AAA71112 | Molecular interact |
| 22 | 18 | 62.1 | 44 | 21 | AAA71125 | Molecular interact |
| 23 | 18 | 62.1 | 44 | 21 | AAA71133 | Molecular interact |
| 24 | 18 | 62.1 | 45 | 21 | AAA70824 | Molecular interact |
| 25 | 18 | 62.1 | 45 | 21 | AAA70825 | Molecular interact |
| 26 | 18 | 62.1 | 45 | 21 | AAA70826 | Molecular interact |
| 27 | 18 | 62.1 | 46 | 21 | AAA71085 | Molecular interact |
| 28 | 18 | 62.1 | 46 | 21 | AAA71087 | Molecular interact |
| 29 | 18 | 62.1 | 46 | 21 | AAA71088 | Molecular interact |
| 30 | 18 | 62.1 | 46 | 21 | AAA71089 | Molecular interact |
| 31 | 18 | 62.1 | 46 | 21 | AAA71090 | Molecular interact |
| 32 | 18 | 62.1 | 46 | 21 | AAA71093 | Molecular interact |
| 33 | 18 | 62.1 | 46 | 21 | AAA71094 | Molecular interact |
| 34 | 18 | 62.1 | 46 | 21 | AAA71095 | Molecular interact |
| 35 | 18 | 62.1 | 46 | 21 | AAA71096 | Molecular interact |
| 36 | 18 | 62.1 | 46 | 21 | AAA71099 | Molecular interact |
| 37 | 18 | 62.1 | 46 | 21 | AAA71100 | Molecular interact |
| 38 | 18 | 62.1 | 46 | 21 | AAA71103 | Molecular interact |
| 39 | 18 | 62.1 | 46 | 21 | AAA71104 | Molecular interact |
| 40 | 18 | 62.1 | 46 | 21 | AAA71105 | Molecular interact |
| 41 | 18 | 62.1 | 46 | 21 | AAA71106 | Molecular interact |
| 42 | 18 | 62.1 | 46 | 21 | AAA71107 | Molecular interact |
| 43 | 18 | 62.1 | 46 | 21 | AAA71109 | Molecular interact |
| 44 | 18 | 62.1 | 46 | 21 | AAA71110 | Molecular interact |
| 45 | 18 | 62.1 | 46 | 21 | AAA71111 | Molecular interact |

ALIGNMENTS

RESULT 1
AAA70827
ID AAA70827 standard; RNA; 29 BP.
XX
AC AAA70827;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #27.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Synthetic.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -
XX

PS Claim 235; Page 235; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 29 BP; 4 A; 4 C; 5 G; 5 U; 11 other;

Query Match 62.1%; Score 18; DB 21; Length 29;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 gauncuuunnguaagcccnangng 27
||||||||||||||||||||||||
Db 4 gauncuuunnguaagcccnangng 27

RESULT 2
AAA70828
ID AAA70828 standard; RNA; 29 BP.
XX
AC AAA70828;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #28.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Homo sapiens.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -
XX

PS Claim 235; Page 235; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 29 BP; 5 A; 5 C; 7 G; 12 U; 0 other;

Query Match 62.1%; Score 18; DB 21; Length 29;
Best Local Similarity 75.0%; Pred. No. 0.35;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 gauncuuunnguaagcccnangng 27
||| |||| |||||||| | | |
Db 4 gauucuuuuuguaagcccuagggg 27

RESULT 3
AAA70829
ID AAA70829 standard; RNA; 29 BP.
XX
AC AAA70829;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #29.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Mus sp.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -
XX

PS Claim 235; Page 235; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 29 BP; 8 A; 6 C; 6 G; 9 U; 0 other;

Query Match 62.1%; Score 18; DB 21; Length 29;
Best Local Similarity 75.0%; Pred. No. 0.35;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 gauncuuunnguaagcccnangng 27
||| |||| |||||||| | | |
Db 4 gauucuuuuuguaagccccaaggg 27

RESULT 4
AAA70830
ID AAA70830 standard; RNA; 29 BP.
XX
AC AAA70830;
XX
DT 27-APR-2001 (first entry)
XX
Molecular interaction site RNA #30.

Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Rattus sp.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -
XX

PS Claim 235; Page 235; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 29 BP; 8 A; 6 C; 6 G; 9 U; 0 other;

Query Match 62.1%; Score 18; DB 21; Length 29;
Best Local Similarity 75.0%; Pred. No. 0.35;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 gauncuuunnguaagcccnangng 27
||| |||| |||||||| | | |
Db 4 gauucuuuuuguaagccccaaggg 27

RESULT 5
AAA71113
ID AAA71113 standard; RNA; 42 BP.
XX
AC AAA71113;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #189.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -
XX

PS Example 7; Figure 122; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 42 BP; 12 A; 7 C; 6 G; 17 U; 0 other;

Query Match 62.1%; Score 18; DB 21; Length 42;
Best Local Similarity 75.0%; Pred. No. 0.37;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 gauncuuunnguaagcccnangng 27
||| |||| |||||||| | | |
Db 7 gauucuuuuuguaagcccuacggg 30

RESULT 6
AAA71114
ID AAA71114 standard; RNA; 42 BP.
XX
AC AAA71114;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #190.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -
XX

PS Example 7; Figure 122; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 42 BP; 11 A; 8 C; 7 G; 16 U; 0 other;

Query Match 62.1%; Score 18; DB 21; Length 42;
Best Local Similarity 75.0%; Pred. No. 0.37;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 gauncuuunnguaagcccnangng 27
||| |||| |||||||| | | |
Db 7 gauucuuuuuguaagcccuaggcg 30

RESULT 7
AAA71115
ID AAA71115 standard; RNA; 42 BP.
XX
AC AAA71115;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #191.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -
XX

PS Example 7; Figure 122; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 42 BP; 13 A; 7 C; 7 G; 15 U; 0 other;

Query Match 62.1%; Score 18; DB 21; Length 42;
Best Local Similarity 75.0%; Pred. No. 0.37;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 gauncuuunnguaagcccnangng 27
||| |||| |||||||| | | |
Db 7 gauucuuuuuguaagccccaaggg 30

RESULT 8
AAA71116
ID AAA71116 standard; RNA; 42 BP.
XX
AC AAA71116;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #192.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -
XX

PS Example 7; Figure 122; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 42 BP; 13 A; 7 C; 7 G; 15 U; 0 other;

Query Match 62.1%; Score 18; DB 21; Length 42;
Best Local Similarity 75.0%; Pred. No. 0.37;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 gauncuuunnguaagcccnangng 27
||| |||| |||||||| | | |
Db 7 gauucuuuuuguaagccccaaggg 30

RESULT 9
AAA71118
ID AAA71118 standard; DNA; 42 BP.
XX
AC AAA71118;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #124.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -
XX

PS Example 7; Figure 125; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 42 BP; 12 A; 7 C; 6 G; 17 T; 0 other;

Query Match 62.1%; Score 18; DB 21; Length 42;
Best Local Similarity 54.2%; Pred. No. 0.37;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

```
Qy      4 gauncuuunnguaagcccnangng 27
          ||: |:::  |:|||||| | | |
Db      7 gattctttttgtaagccctacggg 30
```

RESULT 10
AAA71119
ID AAA71119 standard; DNA; 42 BP.
XX
AC AAA71119;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #125.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -
XX

PS Example 7; Figure 125; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 42 BP; 11 A; 8 C; 7 G; 16 T; 0 other;

Query Match 62.1%; Score 18; DB 21; Length 42;
Best Local Similarity 54.2%; Pred. No. 0.37;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

```
Qy      4 gauncuuunnguaagcccnangng 27
          ||: |:::  |:|||||| | | |
Db      7 gattctttttgtaagccctaggcg 30
```

RESULT 11
AAA71120
ID AAA71120 standard; DNA; 42 BP.
XX
AC AAA71120;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #126.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -
XX

PS Example 7; Figure 125; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 42 BP; 13 A; 7 C; 7 G; 15 T; 0 other;

Query Match 62.1%; Score 18; DB 21; Length 42;
Best Local Similarity 54.2%; Pred. No. 0.37;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 4 gauncuuunnguaagcccnangng 27
|| : |::: |:|||||| | | |
Db 7 gattctttttgtaagccccaaggg 30

RESULT 12
AAA71121
ID AAA71121 standard; DNA; 42 BP.
XX
AC AAA71121;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #127.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -
XX

PS Example 7; Figure 125; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 42 BP; 13 A; 7 C; 7 G; 15 T; 0 other;

Query Match 62.1%; Score 18; DB 21; Length 42;
Best Local Similarity 54.2%; Pred. No. 0.37;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 4 gauncuuunnguaagcccnangng 27
|| : |::: |:|||||| | | |
Db 7 gattctttttgtaagccccaaggg 30

RESULT 13
AAA71123
ID AAA71123 standard; DNA; 42 BP.
XX
AC AAA71123;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #129.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -
XX

PS Example 7; Figure 125; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 42 BP; 9 A; 6 C; 9 G; 18 T; 0 other;

Query Match 62.1%; Score 18; DB 21; Length 42;
Best Local Similarity 54.2%; Pred. No. 0.37;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 4 gauncuuunnguaagcccnangng 27
||: |::: |:|||||| | | |
Db 7 gattctttttgtaagccctagggg 30

RESULT 14
AAA71124
ID AAA71124 standard; DNA; 42 BP.
XX
AC AAA71124;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #130.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -
XX

PS Example 7; Figure 125; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 42 BP; 11 A; 10 C; 7 G; 14 T; 0 other;

Query Match 62.1%; Score 18; DB 21; Length 42;
Best Local Similarity 54.2%; Pred. No. 0.37;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 4 gauncuuunnguaagcccnangng 27
||: |::: |:|||||| | | |
Db 7 gatcctttctgtaagccctacggg 30

RESULT 15
AAA71126
ID AAA71126 standard; RNA; 42 BP.
XX
AC AAA71126;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #195.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -
XX

PS Example 7; Figure 126; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 42 BP; 12 A; 7 C; 6 G; 17 U; 0 other;

Query Match 62.1%; Score 18; DB 21; Length 42;
Best Local Similarity 75.0%; Pred. No. 0.37;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 gauncuuunnguaagcccnangng 27
||| |||| |||||||| | | |
Db 7 gauucuuuuuguaagcccuacggg 30

Search completed: October 2, 2001, 05:03:56
Job time: 3808 sec

GenCore version 4.5


OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 05:00:40 ; Search time 1315.38 Seconds
(without alignments)
341.015 Million cell updates/sec

Title: US-09-310-844B-24
Perfect score: 29
Sequence: 1 uaugauucuuuuuguaagcccuaggggcu 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: em_ba1:*
17: em_ba2:*
18: em_fun:*
19: em_htgo_hum:*
20: em_htgo_inv:*
21: em_htgo_rod:*
22: em_htg_hum1:*
23: em_htg_hum2:*
24: em_htg_hum3:*
25: em_htg_hum4:*
26: em_htg_hum5:*
27: em_htg_hum6:*
28: em_htg_hum7:*
29: em_htg_hum8:*
30: em_htg_inv1:*
31: em_htg_inv2:*
32: em_htg_other:*
33: em_htg_rod:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_om:*
43: em_or:*
44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pl:*
48: em_ro:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_vi:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_vi1:*
59: gb_vi2:*
60: gb_htg1:*
61: gb_htg2:*
62: gb_htg3:*
63: gb_htg4:*
64: gb_htg5:*
65: gb_htg6:*
66: gb_htg7:*
67: gb_htg8:*
68: gb_htg9:*
69: gb_htg10:*
70: gb_htg11:*
71: gb_htg12:*
72: gb_htg13:*
73: gb_htg14:*
74: gb_htg15:*
75: gb_htg16:*
76: gb_htg17:*
77: gb_htg18:*
78: gb_htg19:*
79: gb_htg20:*
80: gb_htg21:*
81: gb_htg22:*
82: gb_htg23:*
83: gb_htg24:*
84: gb_htg25:*
85: gb_pr1:*
86: gb_pr2:*
87: gb_pr3:*
88: gb_pr4:*
89: gb_pr5:*
90: gb_pr6:*
91: gb_pr7:*
92: gb_pr8:*
93: gb_pr9:*
94: gb_ro1:*
95: gb_ro2:*
96: gb_in4:*
97: gb_pr10:*
98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 29 | 100.0 | 209 | 10 | I82323 | I82323 Sequence 8 |
| 2 | 29 | 100.0 | 292 | 54 | G06364 | G06364 human STS W |
| 3 | 29 | 100.0 | 689 | 9 | A17717 | A17717 Plasmid pTG |
| 4 | 29 | 100.0 | 698 | 9 | A02159 | A02159 Synthetic g |
| 5 | 29 | 100.0 | 698 | 9 | A04681 | A04681 Synthetic g |
| 6 | 29 | 100.0 | 698 | 9 | A14089 | A14089 Synthetic D |
| 7 | 29 | 100.0 | 722 | 10 | I03416 | I03416 Sequence 2 |
| 8 | 29 | 100.0 | 730 | 91 | GIBIL2 | K03174 Ape (gibbon |

| | | | | | | |
|---|---|---|---|---|---|---|
| 9 | 29 | 100.0 | 750 | 10 | I01391 | I01391 Sequence 3 |
| 10 | 29 | 100.0 | 769 | 10 | E01460 | E01460 cDNA encodi |
| 11 | 29 | 100.0 | 769 | 10 | E02201 | E02201 DNA encodin |
| 12 | 29 | 100.0 | 782 | 10 | I01101 | I01101 Sequence 2 |
| 13 | 29 | 100.0 | 784 | 9 | A06879 | A06879 Artificial |
| 14 | 29 | 100.0 | 784 | 9 | A14095 | A14095 Synthetic D |
| 15 | 29 | 100.0 | 784 | 9 | A17732 | A17732 Plasmid pTG |
| 16 | 29 | 100.0 | 784 | 10 | E00250 | E00250 DNA coding |
| 17 | 29 | 100.0 | 784 | 10 | I08366 | I08366 Sequence 2 |
| 18 | 29 | 100.0 | 788 | 10 | I07944 | I07944 Sequence 1 |
| 19 | 29 | 100.0 | 790 | 10 | I04498 | I04498 Sequence 1 |
| 20 | 29 | 100.0 | 790 | 10 | I08377 | I08377 Sequence 1 |
| 21 | 29 | 100.0 | 794 | 10 | E00210 | E00210 cDNA encodi |
| 22 | 29 | 100.0 | 794 | 10 | E00211 | E00211 cDNA encodi |
| 23 | 29 | 100.0 | 794 | 10 | E02011 | E02011 cDNA encodi |
| 24 | 29 | 100.0 | 794 | 10 | E02018 | E02018 cDNA encodi |
| 25 | 29 | 100.0 | 795 | 10 | E00214 | E00214 cDNA encodi |
| 26 | 29 | 100.0 | 795 | 10 | E00267 | E00267 DNA sequenc |
| 27 | 29 | 100.0 | 795 | 10 | E00271 | E00271 DNA sequenc |
| 28 | 29 | 100.0 | 801 | 9 | A04961 | A04961 Artificial |
| 29 | 29 | 100.0 | 801 | 9 | A06759 | A06759 Artificial |
| 30 | 29 | 100.0 | 801 | 10 | E00336 | E00336 Human inter |
| 31 | 29 | 100.0 | 801 | 10 | I01197 | I01197 Sequence 1 |
| 32 | 29 | 100.0 | 801 | 10 | I04318 | I04318 Sequence 3 |
| 33 | 29 | 100.0 | 801 | 93 | HSIL02 | V00564 Human mRNA |
| 34 | 29 | 100.0 | 810 | 10 | E00978 | E00978 cDNA sequen |
| 35 | 29 | 100.0 | 810 | 10 | I01387 | I01387 Sequence 1 |
| 36 | 29 | 100.0 | 812 | 10 | E00216 | E00216 DNA encodin |
| 37 | 29 | 100.0 | 812 | 10 | E02540 | E02540 cDNA encodi |
| 38 | 29 | 100.0 | 812 | 97 | S77834 | S77834 Homo sapien |
| 39 | 29 | 100.0 | 825 | 97 | HSU25676 | U25676 Human inter |
| 40 | 29 | 100.0 | 844 | 93 | HSIL2R | X01586 Human mRNA |
| 41 | 29 | 100.0 | 1028 | 97 | S82692 | S82692 interleukin |
| 42 | 29 | 100.0 | 5561 | 97 | HUMIL2B | K02056 Human inter |
| 43 | 29 | 100.0 | 6684 | 93 | HSIL05 | X00695 Human inter |
| 44 | 29 | 100.0 | 6752 | 89 | AF359939 | AF359939 Homo sapi |
| c 45 | 29 | 100.0 | 8491 | 9 | AR031529 | AR031529 Sequence |

ALIGNMENTS

```
RESULT   1
I82323
LOCUS       I82323         209 bp    DNA             PAT       10-JUN-1998
DEFINITION  Sequence 8 from patent US 5712126.
ACCESSION   I82323
VERSION     I82323.1  GI:3210620
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 209)
  AUTHORS   Weissman,S.M. and Prashar,Y.
  TITLE     Analysis of gene expression by display of 3-end restriction
            fragments of CDNA
  JOURNAL   Patent: US 5712126-A 8 27-JAN-1998;
FEATURES             Location/Qualifiers
     source          1. .209
                     /organism="unknown"
BASE COUNT       56 a     30 c     33 g     90 t
ORIGIN


  Query Match           100.0%;  Score 29;  DB 10;  Length 209;
  Best Local Similarity   58.6%;  Pred. No. 0.0038;
  Matches   17;  Conservative   12;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 uaugauucuuuuuguaagcccuaggggcu 29
          :|:||::|:::::|:||||||:||||||:
Db     33 TATGATTCTTTTTGTAAGCCCTAGGGGCT 61
```

```
RESULT   2
G06364
LOCUS       G06364         292 bp    DNA             STS       19-OCT-1995
DEFINITION  human STS WI-7035.
ACCESSION   G06364
VERSION     G06364.1  GI:859609
KEYWORDS    STS sequence; primer; sequence tagged site.
SOURCE      human STSs derived from sequences in dbEST and the Unigene
            collection.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata;
            Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates;
            Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 292)
  AUTHORS   Hudson,T.
  TITLE     Whitehead Institute/MIT Center for Genome Research; Physically
            Mapped ESTs
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Thomas Hudson
            Whitehead Institute/MIT Center for Genome Research
            Whitehead Institute for Biomedical Research
            9 Cambridge Center, Cambridge MA 02142 USA
            Tel: 617 252 1900
            Fax: 617 252 1902
            Email: thudson@genome.wi.mit.edu

            Primer A: TAATTAAGTGCTTCCCACTTAAAAC
            Primer B: ATTTTGGGATAAATAAGTGAAACCA
            STS size: 200
            PCR Profile:
                  Presoak:
                  Denaturation:
                  Annealing: 56 degrees C
                  Polymerization:
                  PCR Cycles: 35
                  Thermal Cycler:
            Protocol:
                Template: 10 ng
                Primer: each 5 pM
                dNTPs: each 4 nM
                Taq Polymerase: 0.025 units/ul
                Total Vol: 20 ul

            Buffer:
                MgCl2: 1.5 mM
                KCl: 50 mM
                Tris-HCL: 10 mM
                pH: 9.3

            Prepared with primer pairs derived from V00564 -- Unigene.
FEATURES             Location/Qualifiers
     source          1. .292
                     /organism="Homo sapiens"
     STS             1. .200
     primer_bind     1. .25
     primer_bind     complement(176. .200)
BASE COUNT       92 a     32 c     32 g    124 t     12 others
ORIGIN


  Query Match           100.0%;  Score 29;  DB 54;  Length 292;
  Best Local Similarity   58.6%;  Pred. No. 0.0039;
  Matches   17;  Conservative   12;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 uaugauucuuuuuguaagcccuaggggcu 29
          :|:||::|:::::|:||||||:||||||:
Db    141 TATGATTCTTTTTGTAAGCCCTAGGGGCT 169
```

```
RESULT   3
A17717
```

```
LOCUS       A17717        689 bp    mRNA            PAT       17-MAY-1994
DEFINITION  Plasmid pTG 26 mRNA for human IL-2.
ACCESSION   A17717
VERSION     A17717.1  GI:513949
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  Plasmid unidentified
            unclassified.
REFERENCE   1  (bases 1 to 689)
  AUTHORS   Kieny,M.P., Sondermeyer,P. and Lecocq,J.P.
  TITLE     Expression of human IL-2 in mammalian cells by a recombined pox
            virus
  JOURNAL   Patent: EP 0206939-A 1 30-DEC-1986;
            TRANSGENE S.A
FEATURES             Location/Qualifiers
     source          1. .689
                     /organism="unidentified"
                     /plasmid="pTG26"
                     /db_xref="taxon:32644"
     CDS             1. .384
                     /partial
                     /codon_start=1
                     /product="human interleukin 2"
                     /protein_id="CAA01347.1"
                     /db_xref="GI:513950"
                     /translation="TKKTQLQLEHLLLDLQMILNGINNYKNPKLTRMLTFKFYMPKKA
                     TELKHLQCLEEELKPLEEVLNLAQSKNFHLRPRDLISNINVIVLELKGSETTFMCEYA
                     DETATIVEFLNRWITFCQSIISTLT"
BASE COUNT      251 a    117 c     99 g    222 t
ORIGIN
```

```
  Query Match             100.0%;  Score 29;  DB 9;  Length 689;
  Best Local Similarity   58.6%;  Pred. No. 0.0043;
  Matches   17;  Conservative   12;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 uaugauucuuuuuguaagcccuaggggcu 29
          :|:||::|:::::|:|||||:||||||:
Db    525 TATGATTCTTTTTGTAAGCCCTAGGGGCT 553
```

```
RESULT   4
A02159
LOCUS       A02159        698 bp    DNA             PAT       21-MAY-1993
DEFINITION  Synthetic gene for interleukin 2 (IL-2) (partial).
ACCESSION   A02159
VERSION     A02159.1  GI:412310
  YWORDS    .
  URCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 698)
  AUTHORS   .
  TITLE     VECTOR FOR THE EXPRESSION IN YEASTS OF INTERLEUKINE-2, TRANSFORMED
            YEASTS AND METHOD FOR PREPARING INTERLEUKINE-2
  JOURNAL   Patent: WO 8503723-A 6 29-AUG-1985;
FEATURES             Location/Qualifiers
     source          1. .698
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     CDS             10. .393
                     /partial
                     /codon_start=1
                     /transl_table=11
                     /product="interleukin 2 (IL-2)"
                     /protein_id="CAA00227.1"
                     /db_xref="GI:412311"
                     /translation="TKKTQLQLEHLLLDLQMILNGINNYKNPKLTRMLTFKFYMPKKA
                     TELKHLQCLEEELKPLEEVLNLAQSKNFHLRPRDLISNINVIVLELKGSETTFMCEYA
                     DETATIVEFLNRWITFCQSIISTLT"
BASE COUNT      251 a    117 c    107 g    222 t      1 others
ORIGIN
```

```
  Query Match             100.0%;  Score 29;  DB 9;  Length 698;
  Best Local Similarity   58.6%;  Pred. No. 0.0043;
  Matches   17;  Conservative   12;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 uaugauucuuuuuguaagcccuaggggcu 29
          :|:||::|:::::|:|||||:||||||:
Db    534 TATGATTCTTTTTGTAAGCCCTAGGGGCT 562
```

```
RESULT   5
A04681
LOCUS       A04681        698 bp    DNA             PAT       24-MAY-1993
DEFINITION  Synthetic gene for interleukin 2 (IL-2) (partial).
ACCESSION   A04681
VERSION     A04681.1  GI:412434
KEYWORDS    interleukin 2.
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 698)
  AUTHORS   Lemoine,Y., Sondermeyer,P., Loison,G., Aigle,M. and Lecocq,J.P.
  TITLE     Yeast-expression vectors for interleukin-2, transformed yeasts and
            process for the preparation of interleukin-2
  JOURNAL   Patent: EP 0152358-A 5 21-AUG-1985;
            TRANSGENE S.A
FEATURES             Location/Qualifiers
     source          1. .698
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     CDS             10. .393
                     /partial
                     /codon_start=1
                     /transl_table=11
                     /product="interleukin 2 (IL-2)"
                     /protein_id="CAA00377.1"
                     /db_xref="GI:412435"
                     /translation="TKKTQLQLEHLLLDLQMILNAINNYKNPKLTRMLTFKFYMPKKA
                     TELKHLQCLEEELKPLEEVLNLAQSKNFHLRPRDLISNINVIVLELKGSETTFMCEYA
                     DETATIVEFLNRWITFCQSIISTLT"
BASE COUNT      251 a    118 c    106 g    222 t      1 others
ORIGIN
```

```
  Query Match             100.0%;  Score 29;  DB 9;  Length 698;
  Best Local Similarity   58.6%;  Pred. No. 0.0043;
  Matches   17;  Conservative   12;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 uaugauucuuuuuguaagcccuagggcu 29
          :|:||::|:::::|:|||||:||||||:
Db    534 TATGATTCTTTTTGTAAGCCCTAGGGCT 562
```

```
RESULT   6
A14089
LOCUS       A14089        698 bp    DNA             PAT       28-FEB-1994
DEFINITION  Synthetic DNA (pTG26) for human interleukin-2 (partial).
ACCESSION   A14089
VERSION     A14089.1  GI:491760
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 698)
  AUTHORS   .
  TITLE     EXPRESSION OF HUMAN IL-2 IN MAMMAL CELLS BY MEANS OF A RECOMBINED
            POXVIRUS
  JOURNAL   Patent: WO 8607610-A 1 31-DEC-1986;
FEATURES             Location/Qualifiers
     source          1. .698
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     CDS             10. .393
```

```
                     /partial
                     /codon_start=1
                     /transl_table=11
                     /product="interleukin 2"
                     /protein_id="CAA01145.1"
                     /db_xref="GI:491761"
                     /translation="TKKTQLQLEHLLLDLQMILNGINNYKNPKLTRMLTFKFYMPKKA
                     TELKHLQCLEEELKPLEEVLNLAQSKNFHLRPRDLISNINVIVLELKGSETTFMCEYA
                     DETATIVEFLNRWITFCQSIISTLT"
BASE COUNT      251 a    117 c    108 g    222 t
ORIGIN


  Query Match             100.0%;  Score 29;  DB 9;  Length 698;
  Best Local Similarity   58.6%;  Pred. No. 0.0043;
  Matches   17;  Conservative   12;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 uaugauucuuuuuguaagcccuaggggcu 29
          :|:||::|:::::|:|||||||:||||||:
Db    534 TATGATTCTTTTTGTAAGCCCTAGGGGCT 562


RESULT   7
I03416
LOCUS       I03416        722 bp ss-DNA             PAT       21-MAY-1993
DEFINITION  Sequence 2 from Patent US 4882282.
ACCESSION   I03416
VERSION     I03416.1  GI:270625
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 722)
  AUTHORS`  Anderson,D.M., Baker,P.E., Cantrell,M.A., Cerretti,D.P.,
            Cosman,D.J., Gimpel,S.D., Grabstein,K.H., Larsen,A.D. and
            McKereghan,K.N.
  TITLE     DNA sequences encoding bovine interleukin-2
  JOURNAL   Patent: US 4882282-A 2 21-NOV-1989;
            Immunex Corporation;
            Seattle, WA
FEATURES             Location/Qualifiers
     source          1. .722
                     /organism="unknown"
BASE COUNT      245 a    130 c    112 g    235 t
ORIGIN


  Query Match             100.0%;  Score 29;  DB 10;  Length 722;
  Best Local Similarity   58.6%;  Pred. No. 0.0043;
  Matches   17;  Conservative   12;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 uaugauucuuuuuguaagcccuaggggcu 29
          :|:||::|:::::|:|||||||:||||||:
Db    610 TATGATTCTTTTTGTAAGCCCTAGGGGCT 638


RESULT   8
GIBIL2
LOCUS       GIBIL2        730 bp    mRNA            PRI       27-APR-1993
DEFINITION  Ape (gibbon) interleukin 2 mRNA.
ACCESSION   K03174
VERSION     K03174.1  GI:177012
KEYWORDS    interleukin; interleukin 2.
SOURCE      Gibbon ('normal gibbon'), cDNA to mRNA, clone CM11.
  ORGANISM  Hylobates lar
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
REFERENCE   1  (bases 1 to 730)
  AUTHORS   Chen,S.J., Holbrook,N.J., Mitchell,K.F., Vallone,C.A.,
            Greengard,J.S., Crabtree,G.R. and Lin,Y.
  TITLE     A viral long terminal repeat in the interleukin 2 gene of a cell
            line that constitutively produces interleukin 2
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 82, 7284-7288 (1985)
  MEDLINE   86042650
COMMENT     [1] also sequenced the IL-2 mRNA from gibbon cell line MLA 144 (see
            separate entry).
FEATURES             Location/Qualifiers
     source          1. .730
                     /organism="Hylobates lar"
                     /db_xref="taxon:9580"
     mRNA            <1. .>729
                     /note="IL-2 mRNA"
     CDS             25. .486
                     /note="interleukin-2"
                     /codon_start=1
                     /protein_id="AAA35453.1"
                     /db_xref="GI:177013"
                     /translation="MYRMQLLSCIALSLALVTNSAPTSSSTKKTQLQLEHLLLDLQMI
                     LNGINNYKNPKLTRMLTFKFYMPKKATELKHLQCLEEELKPLEEVLNLAQSKNFHLRP
                     RDLISNINVIVLELKGSETTFMCEYADETATIVEFLNRWITFCQSIISTLT"
     variation       241
                     /note="c in CM11; t in CM9"
BASE COUNT      247 a    138 c    110 g    235 t
ORIGIN      180 bp upstream of HinfI site.


  Query Match             100.0%;  Score 29;  DB 91;  Length 730;
  Best Local Similarity   58.6%;  Pred. No. 0.0043;
  Matches   17;  Conservative   12;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 uaugauucuuuuuguaagcccuaggggcu 29
          :|:||::|:::::|:|||||||:||||||:
Db    627 TATGATTCTTTTTGTAAGCCCTAGGGGCT 655


RESULT   9
I01391
LOCUS       I01391        750 bp ss-DNA             PAT       21-MAY-1993
DEFINITION  Sequence 3 from Patent US 4851512.
ACCESSION   I01391
VERSION     I01391.1  GI:270228
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 750)
  AUTHORS   Miyaji,H. and Itoh,S.
  TITLE     Novel human interleukin-2 polypeptide derivative
  JOURNAL   Patent: US 4851512-A 3 25-JUL-1989;
            Kyowa Hakko Kogyo Co., Ltd.;
            Tokyo, ;
            JP;
FEATURES             Location/Qualifiers
     source          1. .750
                     /organism="unknown"
BASE COUNT      276 a    131 c    104 g    239 t
ORIGIN


  Query Match             100.0%;  Score 29;  DB 10;  Length 750;
  Best Local Similarity   58.6%;  Pred. No. 0.0044;
  Matches   17;  Conservative   12;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 uaugauucuuuuuguaagcccuaggggcu 29
          :|:||::|:::::|:|||||||:||||||:
Db    588 TATGATTCTTTTTGTAAGCCCTAGGGGCT 616


RESULT  10
E01460
LOCUS       E01460        769 bp    RNA             PAT       29-SEP-1997
DEFINITION  cDNA encoding human interleukin 2.
ACCESSION   E01460
VERSION     E01460.1  GI:2169716
KEYWORDS    JP 1987265992-A/1.
SOURCE      Homo sapiens.
```

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 769)
  AUTHORS   Onomichi,K., Eto,Y. and Shibai,H. .
  TITLE     PRODUCTION OF VALUABLE SUBSTANCE USING HUMAN CELL
  JOURNAL   Patent: JP 1987265992-A 1 18-NOV-1987;
            AJINOMOTO CO INC
COMMENT     OS   Human
            PN   JP 1987265992-A/1
            PD   18-NOV-1987
            PF   12-MAY-1986 JP 1986107981
            PI   ONOMICHI KAZUYA, ETO YUZURU, SHIBAI HIROSHIRO PC
            C12P21/00,C12N15/00,C12P21/02,(C12P21/02,C12R1:91); CC
            strandedness: Double;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No;
            CC   *source: clone=pSDI;
            FH   Key             Location/Qualifiers
            FH
            FT   5'UTR           1. .2
            FT   CDS             3. .507
            FT                   /product='human interleukin 2' FT   3'UTR
            508. .814
            FT   mat_peptide     3. .504
            FT                   /product='human interleukin 2'.
FEATURES              Location/Qualifiers
     source           1. .769
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
BASE COUNT      280 a    132 c    112 g    245 t
ORIGIN


  Query Match             100.0%;  Score 29;  DB 10;  Length 769;
  Best Local Similarity   58.6%;  Pred. No. 0.0044;
  Matches   17;  Conservative   12;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 uaugauucuuuuuguaagcccuaggggcu 29
          :|:||::|:::::|:||||||:||||||:
Db    605 TATGATTCTTTTTGTAAGCCCTAGGGGCT 633


RESULT  11
E02201
LOCUS       E02201        769 bp    RNA             PAT       29-SEP-1997
DEFINITION  DNA encoding interleukin2 (IL-2).
ACCESSION   E02201
VERSION     E02201.1  GI:2170439
KEYWORDS    JP 1990009388-A/2.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 769)
  AUTHORS   Murata,M., Eto,Y. and Shibai,H. .
  TITLE     PRODUCTION OF PHYSIOLOGICALLY ACTIVE PROTEIN
  JOURNAL   Patent: JP 1990009388-A 2 12-JAN-1990;
            AJINOMOTO CO INC
COMMENT     OS   Homo sapiens
            PN   JP 1990009388-A/2
            PD   12-JAN-1990
            PF   08-JUL-1988 JP 1988170142
            PR   09-MAR-1988 JP 88P    55270
            PI   MURATA MASAHIRO,. ETO YUZURU, SHIBAI HIROSHIRO PC
            C12P21/00,C12N15/12,C12N15/26,C12P21/02,(C12P21/02,C12R1:91); CC
            strandedness: Double;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No;
            CC   *source: library=pSD(I)/IL-2;
            CC   Feature is identified by experimental;
            FH   Key             Location/Qualifiers
            FH
            FT   5'UTR           1. .2
            FT   CDS             3. .464
            FT                   /product='interleukin2'
            FT                   /note='IL-2'
            FT   mat_peptide     3. .461
            FT                   /product='interleukin2'
            FT                   /note='IL-2'
            FT   3'UTR           465. .769.
FEATURES              Location/Qualifiers
     source           1. .769
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
BASE COUNT      280 a    132 c    112 g    245 t
ORIGIN


  Query Match             100.0%;  Score 29;  DB 10;  Length 769;
  Best Local Similarity   58.6%;  Pred. No. 0.0044;
  Matches   17;  Conservative   12;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 uaugauucuuuuuguaagcccuaggggcu 29
          :|:||::|:::::|:||||||:||||||:
Db    605 TATGATTCTTTTTGTAAGCCCTAGGGGCT 633


RESULT  12
I01101
LOCUS       I01101        782 bp ss-DNA             PAT       21-MAY-1993
DEFINITION  Sequence 2 from Patent US 4761375.
ACCESSION   I01101
VERSION     I01101.1  GI:269302
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 782)
  AUTHORS   Clark,S.C.
  TITLE     Human interleukin-2 cDNA sequence
  JOURNAL   Patent: US 4761375-A 2 02-AUG-1988;
            Genetics Institute, Inc.;
            Cambridge, MA
FEATURES              Location/Qualifiers
     source           1. .782
                      /organism="unknown"
BASE COUNT      272 a    143 c    113 g    253 t      1 others
ORIGIN


  Query Match             100.0%;  Score 29;  DB 10;  Length 782;
  Best Local Similarity   58.6%;  Pred. No. 0.0044;
  Matches   17;  Conservative   12;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 uaugauucuuuuuguaagcccuaggggcu 29
          :|:||::|:::::|:||||||:||||||:
Db    641 TATGATTCTTTTTGTAAGCCCTAGGGGCT 669


RESULT  13
A06879
LOCUS       A06879        784 bp    DNA             PAT       10-NOV-1993
DEFINITION  Artificial sequence for interleukin-related protein.
ACCESSION   A06879
VERSION     A06879.1  GI:490471
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 784)
  AUTHORS   Fiers,W.C. and Devos,R.R.
  TITLE     DNA sequences, recombinant DNA molecules and processes for
            producing human interleukin two-like polypeptides
```

```
  JOURNAL   Patent: EP 0118977-A 15 19-SEP-1984;
            Biogen, Inc.; BIOGEN, INC
FEATURES             Location/Qualifiers
     source          1. .784
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     CDS             48. .509
                     /codon_start=1
                     /transl_table=11
                     /product="interleukin-related protein"
                     /protein_id="CAA00601.1"
                     /db_xref="GI:490472"
                     /translation="MYRMQLLSCIALSLALVTNSAPTSSSTKKTQLQLEHLLLDLQMI
                    \LNGINNYKNPKLTRMLTFKFYMPKKATELKHLQCLEEELKPLEEVLNLAQSKNFHLRP
                     RDLISNINVIVLELKGSETTFMCEYADETATIVEFLNRWITFCQSIISTLT"
BASE COUNT      267 a    146 c    115 g    256 t
ORIGIN


  Query Match             100.0%;  Score 29;  DB 9;  Length 784;
  Best Local Similarity   58.6%;  Pred. No. 0.0044;
  Matches   17;  Conservative   12;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 uaugauucuuuuuguaagcccuaggggcu 29
           :|:||::|:::::|:||||||:||||||:
Db     650 TATGATTCTTTTTGTAAGCCCTAGGGGCT 678


RESULT  14
A14095
LOCUS       A14095        784 bp    DNA             PAT       28-FEB-1994
DEFINITION  Synthetic DNA (pTG36) for human interleukin-2 (partial).
ACCESSION   A14095
VERSION     A14095.1  GI:491762
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 784)
  AUTHORS   .
  TITLE     EXPRESSION OF HUMAN IL-2 IN MAMMAL CELLS BY MEANS OF A RECOMBINED
            POXVIRUS
  JOURNAL   Patent: WO 8607610-A 7 31-DEC-1986;
FEATURES             Location/Qualifiers
     source          1. .784
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     CDS             13. .474
                     /codon_start=1
                     /transl_table=11
                     /product="interleukin 2"
                     /protein_id="CAA01146.1"
                     /db_xref="GI:491763"
                     /translation="MYRMQLLSCIALSLALVTNSAPTSSSTKKTQLQLEHLLLDLQMI
                     LNGINNYKNPKLTRMLTFKFYMPKKATELKHLQCLEEELKPLEEVLNLAQSKNFHLRP
                     RDLISNINVIVLELKGSETTFMCEYAHETATIVEFLNRWITFCQSIISTLT"
BASE COUNT      273 a    148 c    116 g    247 t
ORIGIN


  Query Match             100.0%;  Score 29;  DB 9;  Length 784;
  Best Local Similarity   58.6%;  Pred. No. 0.0044;
  Matches   17;  Conservative   12;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 uaugauucuuuuuguaagcccuaggggcu 29
           :|:||::|:::::|:||||||:||||||:
Db     615 TATGATTCTTTTTGTAAGCCCTAGGGGCT 643


RESULT  15
A17732
LOCUS       A17732        784 bp    mRNA            PAT       17-MAY-1994
DEFINITION  Plasmid pTG 36 mRNA for IL-2.
ACCESSION   A17732
VERSION     A17732.1  GI:513957
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 784)
  AUTHORS   Kieny,M.P., Sondermeyer,P. and Lecocq,J.P.
  TITLE     Expression of human IL-2 in mammalian cells by a recombined pox
            virus
  JOURNAL   Patent: EP 0206939-A 16 30-DEC-1986;
            TRANSGENE S.A
FEATURES             Location/Qualifiers
     source          1. .784
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     source          13. .784
                     /organism="unidentified"
                     /plasmid="pTG36"
                     /db_xref="taxon:32644"
     CDS             13. .474
                     /codon_start=1
                     /product="interleukin-2"
                     /protein_id="CAA01348.1"
                     /db_xref="GI:513958"
                     /translation="MYRMQLLSCIALSLALVTNSAPTSSSTKKTQLQLEHLLLDLQMI
                     LNGINNYKNPKLTRMLTFKFYMPKKATELKHLQCLEEELKPLEEVLNLAQSKNFHLRP
                     RDLISNINVIVLELKGSETTFMCEYAHETATIVEFLNRWITFCQSIISTLT"
BASE COUNT      273 a    148 c    116 g    247 t
ORIGIN


  Query Match             100.0%;  Score 29;  DB 9;  Length 784;
  Best Local Similarity   58.6%;  Pred. No. 0.0044;
  Matches   17;  Conservative   12;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 uaugauucuuuuuguaagcccuaggggcu 29
           :|:||::|:::::|:||||||:||||||:
Db     615 TATGATTCTTTTTGTAAGCCCTAGGGGCT 643


Search completed: October  2, 2001, 05:00:41
Job time: 5328 sec
```

GenCore version 4.5


OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 05:01:46 ; Search time 57.41 Seconds
(without alignments)
95.628 Million cell updates/sec

Title: US-09-310-844B-24
Perfect score: 29
Sequence: 1 uaugauucuuuuuguaagcccuaggggcu 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 94655562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|---|
| 1 | | 29 | 100.0 | 209 | 1 | US-08-510-032A-8 | Sequence 8, Appli |
| 2 | | 29 | 100.0 | 209 | 3 | US-08-688-514-8 | Sequence 8, Appli |
| 3 | | 29 | 100.0 | 801 | 6 | 5314995-8 | Patent No. 5314995 |
| 4 | c | 29 | 100.0 | 8491 | 2 | US-08-757-439-1 | Sequence 1, Appli |
| 5 | c | 18.4 | 63.4 | 2885 | 2 | US-08-791-849A-15 | Sequence 15, Appl |
| 6 | c | 17.8 | 61.4 | 4093 | 4 | US-07-861-458C-5 | Sequence 5, Appli |
| 7 | | 17 | 58.6 | 8920 | 2 | US-08-446-855A-1 | Sequence 1, Appli |
| 8 | | 17 | 58.6 | 8920 | 4 | US-09-150-741-1 | Sequence 1, Appli |
| 9 | | 16.4 | 56.6 | 512 | 2 | US-08-628-428-4 | Sequence 4, Appli |
| 10 | | 16.4 | 56.6 | 512 | 2 | US-08-628-428-7 | Sequence 7, Appli |
| 11 | | 16.4 | 56.6 | 51259 | 3 | US-08-781-891-209 | Sequence 209, App |
| 12 | c | 16.4 | 56.6 | 51259 | 3 | US-08-781-891-209 | Sequence 209, App |
| 13 | c | 16.2 | 55.9 | 1846 | 3 | US-09-173-581-15 | Sequence 15, Appl |
| 14 | | 16 | 55.2 | 223 | 3 | US-08-621-018B-48 | Sequence 48, Appl |
| 15 | | 16 | 55.2 | 225 | 3 | US-08-621-018B-49 | Sequence 49, Appl |
| 16 | | 16 | 55.2 | 230 | 3 | US-08-621-018B-46 | Sequence 46, Appl |
| 17 | | 16 | 55.2 | 251 | 3 | US-08-621-018B-42 | Sequence 42, Appl |
| 18 | | 16 | 55.2 | 260 | 3 | US-08-621-018B-43 | Sequence 43, Appl |
| 19 | c | 16 | 55.2 | 9493 | 2 | US-08-639-857-23 | Sequence 23, Appl |
| 20 | c | 15.8 | 54.5 | 15202 | 3 | US-08-922-635-21 | Sequence 21, Appl |
| 21 | c | 15.8 | 54.5 | 36519 | 3 | US-08-923-137-2 | Sequence 2, Appli |
| 22 | c | 15.6 | 53.8 | 1166 | 1 | US-08-599-252-95 | Sequence 95, Appl |
| 23 | c | 15.6 | 53.8 | 1166 | 5 | PCT-US96-06352-95 | Sequence 95, Appl |
| 24 | c | 15.6 | 53.8 | 1166 | 5 | PCT-US96-06583-95 | Sequence 95, Appl |
| 25 | | 15.6 | 53.8 | 1801 | 1 | US-08-391-000-41 | Sequence 41, Appl |
| 26 | | 15.6 | 53.8 | 1801 | 2 | US-08-741-931-41 | Sequence 41, Appl |
| 27 | | 15.6 | 53.8 | 1897 | 1 | US-08-245-688-1 | Sequence 1, Appli |
| 28 | | 15.6 | 53.8 | 1897 | 1 | US-08-245-688-3 | Sequence 3, Appli |
| 29 | | 15.6 | 53.8 | 1897 | 1 | US-08-245-688-5 | Sequence 5, Appli |
| 30 | | 15.6 | 53.8 | 1897 | 1 | US-08-245-688-7 | Sequence 7, Appli |
| 31 | | 15.6 | 53.8 | 1897 | 1 | US-08-245-688-9 | Sequence 9, Appli |
| 32 | | 15.6 | 53.8 | 1897 | 1 | US-08-245-688-11 | Sequence 11, Appl |
| 33 | | 15.6 | 53.8 | 2300 | 1 | US-08-103-998-3 | Sequence 3, Appli |
| 34 | c | 15.6 | 53.8 | 2549 | 1 | US-08-470-720-2 | Sequence 2, Appli |
| 35 | | 15.6 | 53.8 | 3611 | 2 | US-08-727-118-1 | Sequence 1, Appli |
| 36 | c | 15.6 | 53.8 | 4964 | 1 | US-08-470-720-5 | Sequence 5, Appli |
| 37 | | 15.6 | 53.8 | 7208 | 3 | US-09-166-186-107 | Sequence 107, App |
| 38 | | 15.6 | 53.8 | 7208 | 4 | US-09-313-932-107 | Sequence 107, App |
| 39 | c | 15.4 | 53.1 | 107 | 1 | US-08-441-591-45 | Sequence 45, Appl |
| 40 | c | 15.4 | 53.1 | 107 | 1 | US-08-303-362A-45 | Sequence 45, Appl |
| 41 | c | 15.4 | 53.1 | 107 | 5 | PCT-US95-05600-62 | Sequence 62, Appl |
| 42 | c | 15.4 | 53.1 | 246 | 3 | US-08-617-860B-14 | Sequence 14, Appl |
| 43 | | 15.4 | 53.1 | 340 | 3 | US-08-441-971-13 | Sequence 13, Appl |
| 44 | | 15.4 | 53.1 | 340 | 4 | US-08-221-653-13 | Sequence 13, Appl |
| 45 | | 15.4 | 53.1 | 340 | 4 | US-08-442-144A-13 | Sequence 13, Appl |

ALIGNMENTS

RESULT 1
US-08-510-032A-8
; Sequence 8, Application US/08510032A
; Patent No. 5712126
; GENERAL INFORMATION:
; APPLICANT: Sherman Weissman and Yarindra Prashar
; TITLE OF INVENTION: Analysis of Gene Expression By Display of 3'-
; TITLE OF INVENTION: end Restriction Fragments of cDNA
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Yahwak & Associates
; STREET: 25 Skytop Drive
; CITY: Trumbull
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06611
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word 4.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/510,032A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: George M. Yahwak
; REGISTRATION NUMBER: 26,824
; REFERENCE/DOCKET NUMBER: Yale
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203)268-1951
; TELEFAX: (203)268-1951
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 209 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-510-032A-8

Query Match 100.0%; Score 29; DB 1; Length 209;
Best Local Similarity 58.6%; Pred. No. 7.4e-05;
Matches 17; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 uaugauucuuuuuguaagcccuaggggcu 29
             :|:||::|:::::|:||||||:||||||:
Db      33 TATGATTCTTTTTGTAAGCCCTAGGGGCT 61
```

```
RESULT   2
US-08-688-514-8
; Sequence 8, Application US/08688514
; Patent No. 6010850
;  GENERAL INFORMATION:
;    APPLICANT:  Sherman Weissman and Yarindra Prashar
;    TITLE OF INVENTION:  Analysis of Gene Expression By Display of 3'-
;    TITLE OF INVENTION:  end Restriction Fragments of cDNA
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Yahwak & Associates
;      STREET:  25 Skytop Drive
;      CITY:  Trumbull
;      STATE:  Connecticut
;      COUNTRY:  USA
;      ZIP:  06611
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  floppy disk
;      COMPUTER:  Macintosh
       OPERATING SYSTEM:  MS-DOS
       SOFTWARE:  Microsoft Word 4.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/688,514
;      FILING DATE:
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  George M. Yahwak
;      REGISTRATION NUMBER:  26,824
;      REFERENCE/DOCKET NUMBER:  Yale
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (203)268-1951
;      TELEFAX:  (203)268-1951
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  209 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-08-688-514-8


  Query Match             100.0%;  Score 29;  DB 3;  Length 209;
  Best Local Similarity   58.6%;  Pred. No. 7.4e-05;
  Matches   17;  Conservative   12;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 uaugauucuuuuuguaagcccuaggggcu 29
           :|:||::|:::::|:||||||:||||||:
         33 TATGATTCTTTTTGTAAGCCCTAGGGGCT 61


RESULT   3
5314995-8
;Patent No. 5314995
;    APPLICANT: FELL, HENRY P.;GAYLE, MARGIT A.
;    TITLE OF INVENTION: THERAPEUTIC INTERLEUKIN-2-ANTIBODY
;BASED FUSION PROTEINS
;    NUMBER OF SEQUENCES: 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/468,390
;      FILING DATE: 22-JAN-1990
;SEQ ID NO:8:
;      LENGTH: 801
5314995-8


  Query Match             100.0%;  Score 29;  DB 6;  Length 801;
  Best Local Similarity   58.6%;  Pred. No. 9.3e-05;
  Matches   17;  Conservative   12;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 uaugauucuuuuuguaagcccuaggggcu 29
           :|:||::|:::::|:||||||:||||||:
Db     650 tatgattctttttgtaagccctaggggct 678


RESULT   4
US-08-757-439-1/c
; Sequence 1, Application US/08757439
; Patent No. 5866371
;  GENERAL INFORMATION:
;    APPLICANT:  BADZIONG, Werner
;    APPLICANT:  HABERMANN, Paul
;    APPLICANT:  MOELLER, Joerg
;    APPLICANT:  ARETZ, Werner
;    TITLE OF INVENTION:  PROCESS FOR USING THE YEAST ADH II
;    TITLE OF INVENTION:  PROMOTER SYSTEM FOR THE PRODUCTION OF HETEROLOGOUS
;    TITLE OF INVENTION:  PROTEINS IN HIGH YIELDS
;    NUMBER OF SEQUENCES:  1
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/757,439
;      FILING DATE:  27-NOV-1996
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE 19544233.4
;      FILING DATE:  28-NOV-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  SANDERCOCK, Colin G.
;      REGISTRATION NUMBER:  31,298
;      REFERENCE/DOCKET NUMBER:  18748/303/HOCE
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8491 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-757-439-1


  Query Match             100.0%;  Score 29;  DB 2;  Length 8491;
  Best Local Similarity   58.6%;  Pred. No. 0.00014;
  Matches   17;  Conservative   12;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 uaugauucuuuuuguaagcccuaggggcu 29
           :|:||::|:::::|:||||||:||||||:
Db    6654 TATGATTCTTTTTGTAAGCCCTAGGGGCT 6626


RESULT   5
US-08-791-849A-15/c
; Sequence 15, Application US/08791849A
; Patent No. 5914449
;  GENERAL INFORMATION:
;    APPLICANT:  Makoto MURASE et al.
;    TITLE OF INVENTION:  Method for Increasing Storage
;    TITLE OF INVENTION:  Lipid Content in Plant Seed
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wenderoth, Lind & Ponack
;      STREET:  805 Fifteenth Street, N.W., #700
```

```
;       CITY:  Washington
;       STATE:  D.C.
;       COUNTRY:  U.S.A.
;       ZIP:  20005
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette, 3.5 inch, 1.44 mb
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  MS-DOS
;       SOFTWARE:  Wordperfect 5.1
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/791,849A
;       FILING DATE:  January 30, 1997
;       CLASSIFICATION:  800
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Warren M. Cheek, Jr.
        REGISTRATION NUMBER:  33,367
        REFERENCE/DOCKET NUMBER:
      TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  202-371-8850
;       TELEFAX:
;       TELEX:
;   INFORMATION FOR SEQ ID NO:  15:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  2885 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     ANTI-SENSE:  NO
;     ORIGINAL SOURCE:
;       ORGANISM:  Saccharomyces cerevisiae
US-08-791-849A-15


  Query Match             63.4%;  Score 18.4;  DB 2;  Length 2885;
  Best Local Similarity   39.3%;  Pred. No. 8.9;
  Matches   11;  Conservative   11;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       2 augauucuuuuuguaagcccuaggggcu 29
           |:||::|::::: :|| |  :|| | |:
Db    2423 ATGATTCTTTTTTTAAACGGTAGAGACT 2396


RESULT   6
US-07-861-458C-5/c
; Sequence 5, Application US/07861458C
; Patent No. 6232061
;  GENERAL INFORMATION:
;     APPLICANT:  Marchionni, Mark Andrew
;     APPLICANT:  Johnson, Carl D.
;     TITLE OF INVENTION:  HOMOLOGY CLONING
;     NUMBER OF SEQUENCES:  142
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Fish & Richardson
;       STREET:  225 Franklin Street
;       CITY:  Boston
;       STATE:  Massachusetts
;       COUNTRY:  U.S.A.
;       ZIP:  02110-2804
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;       COMPUTER:  IBM PS/2 Model 50Z or 55SX
;       OPERATING SYSTEM:  MS-DOS (Version 5.0)
;       SOFTWARE:  WordPerfect (Version 5.1)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/861,458C
;       FILING DATE:  04/01/92
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Clark, Paul T.
;       REGISTRATION NUMBER:  30,162
;       REFERENCE/DOCKET NUMBER:  04585/014001
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (617) 542-5070
;       TELEFAX:  (617) 542-8906
;       TELEX:  200154
;   INFORMATION FOR SEQ ID NO:  5:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  4093
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-07-861-458C-5


  Query Match             61.4%;  Score 17.8;  DB 4;  Length 4093;
  Best Local Similarity   41.4%;  Pred. No. 18;
  Matches   12;  Conservative   10;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 uaugauucuuuuuguaagcccuaggggcu 29
           :|:||::| ::::  ||| | :|| | |:
Db    1146 TATGATTCCTTTTAAAAGTCGTAGTGTCT 1118


RESULT   7
US-08-446-855A-1
; Sequence 1, Application US/08446855A
; Patent No. 5849573
;  GENERAL INFORMATION:
;     APPLICANT:  Stewart, Thomas S
;     APPLICANT:  Flores, Maria V
;     APPLICANT:  O'Sullivan, William J
;     TITLE OF INVENTION:  Nucleotide sequence encoding carbamoyl
;     TITLE OF INVENTION:  phosphate synthetase II
;     NUMBER OF SEQUENCES:  2
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Nixon & Vanderhye PC
;       STREET:  1100 No. 5849573th Glebe Road, 8th Floor
;       CITY:  Arlington
;       STATE:  Virginia
;       COUNTRY:  USA
;       ZIP:  22201-4714
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.24
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/446,855A
;       FILING DATE:  06-Jul-1995
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Mitchard, Leonard C
;       REGISTRATION NUMBER:  29.009
;       REFERENCE/DOCKET NUMBER:  47-80
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  703-816-4000
;       TELEFAX:  703-816-4100
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  8920 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  genomic
US-08-446-855A-1


  Query Match             58.6%;  Score 17;  DB 2;  Length 8920;
  Best Local Similarity   40.0%;  Pred. No. 47;
```

```
  Matches   10;  Conservative   10;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       2 augauucuuuuuguaagcccuaggg 26
           |:||::|::::: :|||   :| ||
Db    2277 ATGATTCTTTTTTTAAGTTATATGG 2301


RESULT   8
US-09-150-741-1
; Sequence 1, Application US/09150741
; Patent No. 6183996
; GENERAL INFORMATION:
;  APPLICANT: Stewart et al.
;  TITLE OF INVENTION: Nucleotide Sequence Encoding Carbamoyl Phosphate
; Patent No. 6183996
;  TITLE OF INVENTION: Synthetase II
;  FILE REFERENCE:
;  CURRENT APPLICATION NUMBER: US/09/150,741
;  CURRENT FILING DATE: 1998-09-10
;  EARLIER APPLICATION NUMBER: PL6380
   EARLIER FILING DATE: 1992-12-16
   EARLIER APPLICATION NUMBER: AU93/00617
   EARLIER FILING DATE: 1993-12-02
;  EARLIER APPLICATION NUMBER: 08/446,855
;  EARLIER FILING DATE: 1995-07-06
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 8920
;   TYPE: DNA
;   ORGANISM: Plasmodium falciparum
US-09-150-741-1


  Query Match             58.6%;  Score 17;  DB 4;  Length 8920;
  Best Local Similarity   40.0%;  Pred. No. 47;
  Matches   10;  Conservative   10;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       2 augauucuuuuuguaagcccuaggg 26
           |:||::|::::: :|||   :| ||
Db    2277 atgattctttttttaagttatatgg 2301


RESULT   9
US-08-628-428-4
; Sequence 4, Application US/08628428
; Patent No. 5885962
;  GENERAL INFORMATION:
     APPLICANT:  Lu, Hsieng
     TITLE OF INVENTION:  SCF ANALOG COMPOSITIONS AND METHODS
     NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amgen Inc.
;      STREET:  1840 DeHavilland Drive
;      CITY:  Thousand Oaks
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  91320-1789
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/628,428
;      FILING DATE:  05-APR-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Knight, Matthew W
;      REGISTRATION NUMBER:  36,846
;      REFERENCE/DOCKET NUMBER:  A-400
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  512 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-628-428-4


  Query Match             56.6%;  Score 16.4;  DB 2;  Length 512;
  Best Local Similarity   42.3%;  Pred. No. 55;
  Matches   11;  Conservative    9;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       4 gauucuuuuuguaagcccuaggggcu 29
           ||:: :::: |:||| |: ||| |:
Db     298 GATTTTTTTAGTAAGTTCTCGGGTCT 323


RESULT  10
US-08-628-428-7
; Sequence 7, Application US/08628428
; Patent No. 5885962
;  GENERAL INFORMATION:
;    APPLICANT:  Lu, Hsieng
;    TITLE OF INVENTION:  SCF ANALOG COMPOSITIONS AND METHODS
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amgen Inc.
;      STREET:  1840 DeHavilland Drive
;      CITY:  Thousand Oaks
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  91320-1789
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/628,428
;      FILING DATE:  05-APR-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Knight, Matthew W
;      REGISTRATION NUMBER:  36,846
;      REFERENCE/DOCKET NUMBER:  A-400
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  512 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-628-428-7


  Query Match             56.6%;  Score 16.4;  DB 2;  Length 512;
  Best Local Similarity   42.3%;  Pred. No. 55;
  Matches   11;  Conservative    9;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       4 gauucuuuuuguaagcccuaggggcu 29
           ||:: :::: |:||| |: ||| |:
Db     298 GATTTTTTTAGTAAGTTCTCGGGTCT 323


RESULT  11
US-08-781-891-209
; Sequence 209, Application US/08781891
; Patent No. 6090620
;  GENERAL INFORMATION:
;    APPLICANT:  Fu, Ying-Hui
;    APPLICANT:  Yu, Chang-En
;    APPLICANT:  Oshima, Junko
;    APPLICANT:  Mulligan, John T.
```

```
;    APPLICANT:  Schellenberg, Gerald D.
;    TITLE OF INVENTION:  GENE AND GENE PRODUCTS RELATED TO
;    TITLE OF INVENTION:  WERNER'S SYNDROME
;    NUMBER OF SEQUENCES:  209
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/781,891
;      FILING DATE:  27-DEC-1996
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 6090620tenburg Ph.D., Carol
;      REGISTRATION NUMBER:  39,317
;      REFERENCE/DOCKET NUMBER:  240052.419
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  209:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  51259 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-781-891-209


  Query Match             56.6%;  Score 16.4;  DB 3;  Length 51259;
  Best Local Similarity   42.3%;  Pred. No. 1.2e+02;
  Matches   11;  Conservative    9;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      4 gauucuuuuuguaagcccuaggggcu 29
          | ::|:::::| |||| |:  | ||:
Db  26016 GTTTCTTTTTGGAAGCACTTAGAGCT 26041


RESULT  12
US-08-781-891-209/c
; Sequence 209, Application US/08781891
; Patent No. 6090620
;  GENERAL INFORMATION:
;    APPLICANT:  Fu, Ying-Hui
;    APPLICANT:  Yu, Chang-En
;    APPLICANT:  Oshima, Junko
;    APPLICANT:  Mulligan, John T.
;    APPLICANT:  Schellenberg, Gerald D.
;    TITLE OF INVENTION:  GENE AND GENE PRODUCTS RELATED TO
;    TITLE OF INVENTION:  WERNER'S SYNDROME
;    NUMBER OF SEQUENCES:  209
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/781,891
;      FILING DATE:  27-DEC-1996
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 6090620tenburg Ph.D., Carol
;      REGISTRATION NUMBER:  39,317
;      REFERENCE/DOCKET NUMBER:  240052.419
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  209:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  51259 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-781-891-209


  Query Match             56.6%;  Score 16.4;  DB 3;  Length 51259;
  Best Local Similarity   42.3%;  Pred. No. 1.2e+02;
  Matches   11;  Conservative    9;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      4 gauucuuuuuguaagcccuaggggcu 29
          | ::|:::::| |||| |:  | ||:
Db  10495 GTTTCTTTTTGGAAGCACTTAGAGCT 10470


RESULT  13
US-09-173-581-15/c
; Sequence 15, Application US/09173581A
; Patent No. 6013455
; GENERAL INFORMATION:
;  APPLICANT: Bandman, Olga
;  APPLICANT: Tang, Y. Tom
;  APPLICANT: Hillman, Jennifer L.
;  APPLICANT: Yue, Henry
;  APPLICANT: Guegler, Karl J.
;  APPLICANT: Corley, Neil C.
;  APPLICANT: Gorgone, Gina
;  APPLICANT: Azimzai, Yalda
;  APPLICANT: Lu, Aina
;  TITLE OF INVENTION: Protein Kinase Homologs
;  FILE REFERENCE: PF-0614 US
;  CURRENT APPLICATION NUMBER: US/09/173,581A
;  CURRENT FILING DATE: 1998-10-15
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: PERL Program
; SEQ ID NO 15
;   LENGTH: 1846
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE: -
;   OTHER INFORMATION: 1567782
US-09-173-581-15


  Query Match             55.9%;  Score 16.2;  DB 3;  Length 1846;
  Best Local Similarity   38.1%;  Pred. No. 85;
  Matches    8;  Conservative   10;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 augauucuuuuuguaagcccu 22
          |:||::|::::: :||| | :
Db    155 ATGATTCTTTTTCTAAGGCAT 135


RESULT  14
US-08-621-018B-48
; Sequence 48, Application US/08621018B
; Patent No. 6060239
;  GENERAL INFORMATION:
;    APPLICANT:  Stuart, Susan G.
;    APPLICANT:  Hawkins, Phillip R.
;    APPLICANT:  Seilhamer, Jeffrey J.
;    APPLICANT:  Murry, Lynn E.
```

```
;    TITLE OF INVENTION:  CELLUBREVIN HOMOLOGS
;    NUMBER OF SEQUENCES:  51
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  U.S.
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,018B
;      FILING DATE:  March 22, 1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/409,373
;      FILING DATE:  March 23, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CERRONE, MICHAEL C.
;      REGISTRATION NUMBER:  39,132
;      REFERENCE/DOCKET NUMBER:  PF-0029-1 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;  INFORMATION FOR SEQ ID NO:  48:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  223 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    IMMEDIATE SOURCE:
;      LIBRARY:  LUNGNOT02
;      CLONE:  375070
US-08-621-018B-48


  Query Match             55.2%;  Score 16;  DB 3;  Length 223;
  Best Local Similarity   41.7%;  Pred. No. 73;
  Matches   10;  Conservative    9;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      6 uucuuuuuguaagcccuaggggcu 29
          : |:::::|:|||| |:|  | |:
Db    122 TACTTTTTGTAAGCACTACTGACT 145


SULT  15
S-08-621-018B-49
; Sequence 49, Application US/08621018B
; Patent No. 6060239
;  GENERAL INFORMATION:
;    APPLICANT:  Stuart, Susan G.
;    APPLICANT:  Hawkins, Phillip R.
;    APPLICANT:  Seilhamer, Jeffrey J.
;    APPLICANT:  Murry, Lynn E.
;    TITLE OF INVENTION:  CELLUBREVIN HOMOLOGS
;    NUMBER OF SEQUENCES:  51
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  U.S.
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,018B
;      FILING DATE:  March 22, 1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/409,373
;      FILING DATE:  March 23, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CERRONE, MICHAEL C.
;      REGISTRATION NUMBER:  39,132
;      REFERENCE/DOCKET NUMBER:  PF-0029-1 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;  INFORMATION FOR SEQ ID NO:  49:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  225 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    IMMEDIATE SOURCE:
;      LIBRARY:  LATRNOT01
;      CLONE:  465647
US-08-621-018B-49


  Query Match             55.2%;  Score 16;  DB 3;  Length 225;
  Best Local Similarity   41.7%;  Pred. No. 74;
  Matches   10;  Conservative    9;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      6 uucuuuuuguaagcccuaggggcu 29
          : |:::::|:|||| |:|  | |:
Db     70 TACTTTTTGTAAGCACTACTGACT 93


Search completed: October  2, 2001, 05:01:49
Job time: 5316 sec
```

GenCore version 4.5


OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 04:38:38 ; Search time 1643.83 Seconds
(without alignments)
166.765 Million cell updates/sec

Title: US-09-310-844B-24
Perfect score: 29
Sequence: 1 uaugauucuuuuuguaagcccuaggggcu 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est33:*
26: gb_est34:*
27: gb_est35:*
28: gb_est36:*
29: gb_est37:*
30: gb_est38:*
31: gb_est39:*
32: gb_est40:*
33: em_estba:*
34: em_estfun:*
35: em_esthum1:*
36: em_esthum2:*
37: em_esthum3:*
38: em_esthum4:*
39: em_esthum5:*
40: em_esthum6:*
41: em_esthum7:*
42: em_esthum8:*
43: em_esthum9:*
44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estin5:*
68: em_estom1:*
69: em_estom2:*
70: em_estov1:*
71: em_estov2:*
72: em_estpl1:*
73: em_estpl2:*
74: em_estpl3:*
75: em_estpl4:*
76: em_estpl5:*
77: em_estpl6:*
78: em_estpl7:*
79: em_estpl8:*
80: em_estpl9:*
81: em_estpl10:*
82: em_estro1:*
83: em_estro2:*
84: em_estro3:*
85: em_estro4:*
86: em_estro5:*
87: em_estro6:*
88: em_estro7:*
89: em_estro8:*
90: em_estro9:*
91: em_estro10:*
92: em_estro11:*
93: em_estro12:*
94: em_estro13:*
95: em_estro14:*
96: em_estro15:*
97: em_estro16:*
98: em_estro17:*
99: em_estro18:*
100: em_estro19:*
101: em_estro20:*
102: gb_est25:*
103: gb_est26:*
104: gb_est27:*
105: gb_est28:*
106: gb_est29:*
107: gb_est30:*
108: gb_est31:*
109: gb_est32:*
110: gb_est41:*
111: gb_est42:*
112: gb_est43:*
113: gb_est44:*
114: gb_est45:*
115: gb_est46:*
116: gb_est47:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est75:*
137: gb_est76:*
138: gb_est77:*
139: gb_est78:*
140: gb_est79:*
141: gb_est80:*
142: gb_est81:*
143: gb_est82:*
144: gb_est83:*
145: gb_est84:*
146: gb_est85:*
147: gb_est86:*
148: gb_est87:*
149: gb_est88:*
150: gb_est89:*
151: gb_est90:*
152: gb_est99:*
153: gb_est100:*
154: gb_est101:*
155: gb_est102:*
156: gb_est103:*
157: gb_est104:*
158: gb_est105:*
159: gb_est106:*
160: gb_est67:*
161: gb_est68:*
162: gb_est69:*
163: gb_est70:*
164: gb_est71:*
165: gb_est72:*
166: gb_est73:*
167: gb_est74:*
168: gb_est91:*
169: gb_est92:*
170: gb_est93:*
171: gb_est94:*
172: gb_est95:*
173: gb_est96:*
174: gb_est97:*
175: gb_est98:*
176: em_esthum29:*
177: em_esthum30:*
178: em_esthum31:*
179: em_esthum32:*
180: em_esthum33:*
181: em_estom3:*
182: em_estpl11:*
183: em_estro21:*
184: em_estro22:*
185: em_estro23:*
186: em_htc:*
187: gb_est107:*
188: gb_est108:*
189: gb_est109:*
190: gb_est110:*
191: gb_est111:*
192: gb_htc:*
193: em_gss_fun:*
194: em_gss_hum1:*
195: em_gss_hum2:*
196: em_gss_hum3:*
197: em_gss_hum4:*
198: em_gss_hum5:*
199: em_gss_hum6:*
200: em_gss_hum7:*
201: em_gss_hum8:*
202: em_gss_hum9:*
203: em_gss_inv1:*
204: em_gss_inv2:*
205: em_gss_inv3:*
206: em_gss_other:*
207: em_gss_pln1:*
208: em_gss_pln2:*
209: em_gss_pro:*
210: em_gss_rod1:*
211: em_gss_rod2:*
212: em_gss_rod3:*
213: em_gss_rod4:*
214: em_gss_rod5:*
215: em_gss_vrt1:*
216: em_gss_vrt2:*
217: em_gss_vrt3:*
218: gb_gss1:*
219: gb_gss2:*
220: gb_gss3:*
221: gb_gss4:*
222: gb_gss5:*
223: gb_gss6:*
224: gb_gss7:*
225: gb_gss8:*
226: gb_gss9:*
227: gb_gss10:*
228: gb_gss11:*
229: gb_gss12:*
230: gb_gss13:*
231: gb_gss14:*
232: gb_gss15:*
233: gb_gss16:*
234: gb_gss17:*
235: gb_gss18:*
236: gb_gss19:*
237: gb_gss20:*
238: gb_gss21:*
239: gb_gss22:*
240: gb_gss23:*
241: gb_gss24:*
242: gb_gss25:*
243: gb_gss26:*
244: gb_gss27:*
245: gb_gss28:*
246: gb_gss29:*
247: gb_gss30:*
248: gb_gss31:*
249: gb_gss32:*
250: gb_gss33:*
251: gb_gss34:*
252: em_gss_inv4:*
253: em_gss_rod6:*
254: em_gss_rod7:*
255: em_gss_rod8:*
256: gb_gss35:*
257: gb_gss36:*
258: gb_gss37:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| c 1 | 29 | 100.0 | 527 | 226 | AQ267968 | AQ267968 RPCI11-73 |
| c 2 | 29 | 100.0 | 596 | 226 | AQ267210 | AQ267210 RPCI11-73 |
| 3 | 21.2 | 73.1 | 325 | 4 | AA290106 | AA290106 vb32e03.r |
| c 4 | 20.6 | 71.0 | 914 | 245 | AZ544446 | AZ544446 ENTDF41TF |
| 5 | 19.6 | 67.6 | 494 | 113 | AW240011 | AW240011 ptr1c.pk0 |
| 6 | 19.6 | 67.6 | 619 | 242 | AZ349700 | AZ349700 1M0086K17 |
| 7 | 19.6 | 67.6 | 741 | 251 | AZ899124 | AZ899124 RPCI-24-2 |
| c 8 | 19.6 | 67.6 | 840 | 248 | AZ721454 | AZ721454 RPCI-24-1 |
| c 9 | 19.4 | 66.9 | 189 | 161 | BB588162 | BB588162 BB588162 |
| 10 | 19.4 | 66.9 | 367 | 122 | AW939343 | AW939343 QV1-DT007 |
| 11 | 19.4 | 66.9 | 666 | 246 | AZ579747 | AZ579747 1M0367H06 |
| c 12 | 19.2 | 66.2 | 309 | 226 | AQ239971 | AQ239971 CIT-HSP-2 |
| c 13 | 19 | 65.5 | 348 | 224 | AQ072238 | AQ072238 HS_3024_B |
| 14 | 19 | 65.5 | 358 | 102 | AI808918 | AI808918 wf66h01.x |
| 15 | 19 | 65.5 | 408 | 15 | AI047736 | AI047736 uh82f05.r |
| 16 | 19 | 65.5 | 493 | 10 | AA672498 | AA672498 vo59c04.r |
| c 17 | 19 | 65.5 | 497 | 226 | AQ308567 | AQ308567 CITBI-E1- |
| c 18 | 19 | 65.5 | 521 | 110 | AW045653 | AW045653 UI-M-BH1- |
| c 19 | 19 | 65.5 | 552 | 227 | AQ346030 | AQ346030 RPCI11-13 |
| 20 | 19 | 65.5 | 828 | 151 | BF625129 | BF625129 HVSMEa000 |
| 21 | 18.8 | 64.8 | 86 | 220 | CNS02IOD | AL199174 Tetraodon |
| c 22 | 18.8 | 64.8 | 544 | 170 | BF889394 | BF889394 PM3-TN010 |
| c 23 | 18.8 | 64.8 | 859 | 219 | CNS0112P | AL099739 Drosophil |
| 24 | 18.8 | 64.8 | 1101 | 219 | CNS00ECJ | AL068919 Drosophil |
| c 25 | 18.6 | 64.1 | 173 | 238 | AZ118531 | AZ118531 RPCI-23-1 |
| 26 | 18.6 | 64.1 | 340 | 116 | AW446751 | AW446751 87173 MAR |
| c 27 | 18.6 | 64.1 | 358 | 172 | BG003653 | BG003653 QV4-GN012 |
| c 28 | 18.6 | 64.1 | 605 | 228 | AQ419605 | AQ419605 RPCI-11-1 |
| 29 | 18.6 | 64.1 | 629 | 235 | AQ874895 | AQ874895 V118C5 mT |
| 30 | 18.6 | 64.1 | 688 | 145 | BF168971 | BF168971 601775241 |
| c 31 | 18.6 | 64.1 | 741 | 245 | AZ519681 | AZ519681 RPCI-11-2 |
| c 32 | 18.4 | 63.4 | 235 | 103 | AI875797 | AI875797 uj19h11.x |
| 33 | 18.4 | 63.4 | 287 | 135 | BB515371 | BB515371 BB515371 |
| c 34 | 18.4 | 63.4 | 326 | 118 | AW620862 | AW620862 sj47f04.y |
| c 35 | 18.4 | 63.4 | 338 | 162 | BE020565 | BE020565 sm45c05.y |
| 36 | 18.4 | 63.4 | 352 | 188 | T38192 | T38192 EST103590 S |
| c 37 | 18.4 | 63.4 | 355 | 6 | AA397709 | AA397709 zt87f03.r |
| c 38 | 18.4 | 63.4 | 382 | 20 | AI443236 | AI443236 sa46b06.y |
| 39 | 18.4 | 63.4 | 383 | 7 | AA399048 | AA399048 zt87f01.s |
| 40 | 18.4 | 63.4 | 389 | 233 | AQ790571 | AQ790571 HS_3120_B |
| c 41 | 18.4 | 63.4 | 424 | 154 | BG509674 | BG509674 sad20a12. |
| c 42 | 18.4 | 63.4 | 428 | 151 | BF596933 | BF596933 su82e04.y |
| c 43 | 18.4 | 63.4 | 450 | 21 | AI548206 | AI548206 UI-R-C3-s |
| c 44 | 18.4 | 63.4 | 457 | 116 | AW457912 | AW457912 sh97b08.y |
| c 45 | 18.4 | 63.4 | 470 | 172 | BG042345 | BG042345 sv06c11.y |

ALIGNMENTS

RESULT 1
AQ267968/c
LOCUS AQ267968 527 bp DNA GSS 27-APR-1999
DEFINITION RPCI11-73E11.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-73E11, DNA sequence.
ACCESSION AQ267968
VERSION AQ267968.1 GI:3795572
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 527)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
TITLE Use of human BAC End Sequences for Sequence-Ready Map Building
JOURNAL Unpublished (1998)
COMMENT Other_GSSs: RPCI11-73E11.TK
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.
FEATURES Location/Qualifiers
source 1. .527
/organism="Homo sapiens"
/db_xref="GDB:7527754"
/db_xref="taxon:9606"
/clone="RPCI-11-73E11"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; RPCI11 Human Male BAC Library"
BASE COUNT 203 a 78 c 77 g 167 t 2 others
ORIGIN

Query Match 100.0%; Score 29; DB 226; Length 527;
Best Local Similarity 58.6%; Pred. No. 0.014;
Matches 17; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 uaugauucuuuuuguaagcccuaggggcu 29
           :|:||::|:::::|:||||||:||||||:
Db     358 TATGATTCTTTTTGTAAGCCCTAGGGGCT 330
```

RESULT 2
AQ267210/c
LOCUS AQ267210 596 bp DNA GSS 27-APR-1999
DEFINITION RPCI11-73C7.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-73C7, DNA sequence.
ACCESSION AQ267210
VERSION AQ267210.1 GI:3794814
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 596)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
TITLE Use of human BAC End Sequences for Sequence-Ready Map Building
JOURNAL Unpublished (1998)
COMMENT Other_GSSs: RPCI11-73C7.TK
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.

```
FEATURES             Location/Qualifiers
     source          1. .596
                     /organism="Homo sapiens"
                     /db_xref="GDB:7527702"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-73C7"
                     /clone_lib="RPCI-11"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
BASE COUNT      227 a     92 c     88 g    186 t      3 others
ORIGIN


  Query Match             100.0%;  Score 29;  DB 226;  Length 596;
  Best Local Similarity   58.6%;  Pred. No. 0.015;
  Matches   17;  Conservative   12;  Mismatches    0;  Indels    0;  Gaps    0;

      1 uaugauucuuuuuguaagcccuaggggcu 29
        :|:||::|:::::|:||||||:||||||:
    358 TATGATTCTTTTTGTAAGCCCTAGGGGCT 330


RESULT   3
AA290106
LOCUS       AA290106      325 bp    mRNA            EST       14-APR-1997
DEFINITION  vb32e03.r1 Soares mouse lymph node NbMLN Mus musculus cDNA clone
            IMAGE:750652 5' similar to gb:M16762 Mouse interleukin 2 (MOUSE);,
            mRNA sequence.
ACCESSION   AA290106
VERSION     AA290106.1  GI:1936335
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 325)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:459636
            Putative full length read
            vector to vector length is 422
            Seq primer: -28m13 rev2 ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .325
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:750652"
                     /clone_lib="Soares mouse lymph node NbMLN"
                     /sex="male"
                     /tissue_type="lymph node"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
                     /note="Organ: lymph node; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Not I - oligo(dT) primer
                     [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGATACTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. RNA
                     provided by Dr. Bertrand Jordan. Library constructed and
                     normalized by Bento Soares and M.Fatima Bonaldo."
BASE COUNT      107 a     47 c     41 g    130 t
ORIGIN


  Query Match             73.1%;  Score 21.2;  DB 4;  Length 325;
  Best Local Similarity   57.7%;  Pred. No. 38;
  Matches   15;  Conservative    8;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 augauucuuuuuguaagcccuagggg 27
           | ||::|:::::|:|||||| | |||
Db     168 AAGATTCTTTTTGTAAGCCCCAAGGG 193


RESULT   4
AZ544446/c
LOCUS       AZ544446      914 bp    DNA             GSS       14-NOV-2000
DEFINITION  ENTDF41TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
            genomic, DNA sequence.
ACCESSION   AZ544446
VERSION     AZ544446.1  GI:11164291
KEYWORDS    GSS.
SOURCE      Entamoeba histolytica.
  ORGANISM  Entamoeba histolytica
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1  (bases 1 to 914)
  AUTHORS   Loftus,B., Van Aken,S. and Fraser,C.
  TITLE     Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: bjloftus@tigr.org
            Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
            DNA library
            Seq primer: M13-Forward
            Class: shotgun
            High quality sequence start: 96
            High quality sequence stop: 775.
FEATURES             Location/Qualifiers
     source          1. .914
                     /organism="Entamoeba histolytica"
                     /strain="HM1:IMSS"
                     /db_xref="taxon:5759"
                     /clone_lib="Entamoeba histolytica Sheared DNA"
                     /note="Vector: pHOS1; Site_1: Bst I; Constructed at The
                     Institute for Genomic Research (TIGR), Rockville, MD.
                     Genomic DNA isolated from broth cultures of E. histolytica
                     using a  method described by Clark and Diamond (Clark,
                     C.G., and Diamond, L.S.  (1993) Entamoeba histolytica: a
                     method for isolate identification.  Exp. Parasitol.
                     77:450.). The DNA was mechanically sheared to give a
                     tight size distribution (~2 kb). The v + i method used for
                     the library  construction is described in detail in Smith,
                     H.O. and Venter, J.C.  (Making small insert libraries for
                     whole genome shotgun sequencing  projects. In Genome
                     Sequencing: A Practical Approach, eds. M. Vaudin  and B.
                     Barell, Oxford University Press, 1999)."
BASE COUNT      231 a    238 c    208 g    237 t
ORIGIN


  Query Match             71.0%;  Score 20.6;  DB 245;  Length 914;
  Best Local Similarity   48.1%;  Pred. No. 81;
```

```
  Matches   13;  Conservative   10;  Mismatches   4;  Indels    0;  Gaps    0;

Qy      3 ugauucuuuuuguaagcccuaggggcu 29
          :|| : :::::|:|||||||:  ||||:
Db    394 TGACTTTTTTTGTAAGCCCTGCGGGCT 368


RESULT   5
AW240011
LOCUS       AW240011      494 bp    mRNA            EST       27-APR-2000
DEFINITION  ptr1c.pk001.120 chicken T cell cDNA library Gallus gallus cDNA
            clone ptr1c.pk001.120 5', mRNA sequence.
ACCESSION   AW240011
VERSION     AW240011.1  GI:6579751
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
  FERENCE   1  (bases 1 to 494)
  AUTHORS   Morgan,R.
  TITLE     Chicken T cell ESTs
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robin Morgan
            University of Delaware
            Townsend Hall, Newark, DE 19717, USA
            Tel: 302-831-1341
            Fax: 302-831-2822
            Email: morgan@udel.edu
            Clones can be ordered online at http://www.chickest.udel.edu.
FEATURES             Location/Qualifiers
     source          1. .494
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
                     /clone="ptr1c.pk001.120"
                     /clone_lib="chicken T cell cDNA library"
                     /sex="Male and Female"
                     /cell_type="splenic T cell"
                     /lab_host="E.coli DH10B"
                     /note="Vector: pB42AD"
BASE COUNT      151 a     82 c     96 g    143 t     22 others
ORIGIN


  Query Match             67.6%;  Score 19.6;  DB 113;  Length 494;
  Best Local Similarity   46.4%;  Pred. No. 2.1e+02;
  Matches   13;  Conservative    9;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      2 augauucuuuuuguaagcccuaggggcu 29
          |:||: |:::::|:|||||| |    |:
Db     16 ATGATGCTTTTTGTAAGCCCNANCTACT 43


RESULT   6
AZ349700
LOCUS       AZ349700      619 bp    DNA             GSS       29-SEP-2000
DEFINITION  1M0086K17R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0086K17 R, DNA sequence.
ACCESSION   AZ349700
VERSION     AZ349700.1  GI:10428937
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 619)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0086  row: K  column: 17
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 619.
FEATURES             Location/Qualifiers
     source          1. .619
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0086K17"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT      111 a    173 c    106 g    229 t
ORIGIN


  Query Match             67.6%;  Score 19.6;  DB 242;  Length 619;
  Best Local Similarity   53.8%;  Pred. No. 2.1e+02;
  Matches   14;  Conservative    8;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      4 gauucuuuuuguaagcccuaggggcu 29
          ||::|::: :|  |||||:||| ||:
Db    289 GATTCTTTCTGCGAGCCCTAGGTGCT 314


RESULT   7
AZ899124
LOCUS       AZ899124      741 bp    DNA             GSS       05-MAR-2001
DEFINITION  RPCI-24-211B11.TV RPCI-24 Mus musculus genomic clone RPCI-24-211B11
            , DNA sequence.
ACCESSION   AZ899124
VERSION     AZ899124.1  GI:13218069
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 741)
  AUTHORS   Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M.,
            Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E.,
            Russell,D., de Jong,P. and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-24
  JOURNAL   Unpublished (1999)
```

```
COMMENT     Other_GSSs: RPCI-24-211B11.TJ
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-24. For BAC
            library availability, please contact Pieter de Jong
            (pdejong@mail.cho.org). Clones may be purchased from BACPAC
            Resources (http://www.chori.org/bacpac/orderingframe.htm). BAC end
            page: http://ww .tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 211  row: B  column: 11
            Seq primer: T7
            Class: BAC ends.
FEATURES              Location/Qualifiers
     source           1. .741
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="RPCI-24-211B11"
                      /clone_lib="RPCI-24"
                      /sex="Male"
                      /cell_type="Spleen/Brain"
                      /note="Vector: pTARBAC1; Site_1: BamH1; Site_2: BamH1;
                      RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
                      library was  cloned in the pTARBAC1 cloning vector at the
                      BamH1 sites using MboI  partially digested male C57BL/6J
                      DNA."
BASE COUNT      199 a    174 c    151 g    217 t
ORIGIN


  Query Match             67.6%;  Score 19.6;  DB 251;  Length 741;
  Best Local Similarity   46.2%;  Pred. No. 2.2e+02;
  Matches   12;  Conservative   10;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 uaugauucuuuuuguaagcccuaggg 26
          :|: |::|::: :|:| ||||: |||
Db    701 TATTATTCTTTCTGTATGCCCTGGGG 726


RESULT   8
AZ721454/c
LOCUS       AZ721454      840 bp    DNA             GSS       24-JAN-2001
DEFINITION  RPCI-24-144G10.TJ RPCI-24 Mus musculus genomic clone RPCI-24-144G10
            , DNA sequence.
  CESSION   AZ721454
  RSION     AZ721454.1  GI:12464163
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 840)
  AUTHORS   Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M.,
            Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E.,
            Russell,D., de Jong,P. and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-24
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: RPCI-24-144G10.TV
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-24. For BAC
            library availability, please contact Pieter de Jong
            (pdejong@mail.cho.org). Clones may be purchased from BACPAC
            Resources (http://www.chori.org/bacpac/orderingframe.htm). BAC end
            page: http://ww .tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 144  row: G  column: 10
            Seq primer: SP6
            Class: BAC ends.
FEATURES              Location/Qualifiers
     source           1. .840
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="RPCI-24-144G10"
                      /clone_lib="RPCI-24"
                      /sex="Male"
                      /cell_type="Spleen/Brain"
                      /note="Vector: pTARBAC1; Site_1: BamH1; Site_2: BamH1;
                      RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
                      library was  cloned in the pTARBAC1 cloning vector at the
                      BamH1 sites using MboI  partially digested male C57BL/6J
                      DNA."
BASE COUNT      365 a    173 c    136 g    166 t
ORIGIN


  Query Match             67.6%;  Score 19.6;  DB 248;  Length 840;
  Best Local Similarity   46.2%;  Pred. No. 2.2e+02;
  Matches   12;  Conservative   10;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      4 gauucuuuuuguaagcccuaggggcu 29
          | ::|:::::|:| || |: |||||:
Db    440 GTTTCTTTTTGTAGGCACTCGGGGCT 415


RESULT   9
BB588162/c
LOCUS       BB588162      189 bp    mRNA            EST       30-NOV-2000
DEFINITION  BB588162 RIKEN full-length enriched, 6 days neonate skin Mus
            musculus cDNA clone A030005O22 5', mRNA sequence.
ACCESSION   BB588162
VERSION     BB588162.1  GI:11484706
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 189)
  AUTHORS   Aizawa,K., Akahira,S., Akimura,T., Arai,A., Arakawa,T., Carninci,P.
            , Hanagaki,T., Hayatsu,N., Hiraoka,T., Hirozane,T., Hodoyama,Y.,
            Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kawai,J., Kojima,Y., Konno
            ,H., Kusakabe,M., Matsuyama,T., Miyazaki,A., Nakamura,M., Nishi,K.,
            Nomura,K., Numazaki,R., Okazaki,Y., Okido,T., Owa,C., Sakai,C.,
            Sakai,K., Sasaki,D., Sato,K., Shibata,K., Shibata,Y., Shinagawa,A.,
            Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Tanaka
            ,T., Toya,T., Watahiki,A., Yamamura,T., Yasunishi,A., Yoshida,K.,
            Yoshiki,A., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Aizawa,K. et al. 2000)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Yoshihide Hayashizaki
            Genome Exploration Research Group, Life Science Tsukuba Center,
            Genome Science Laboratory
            The Institute of Physical and Chemical Research (RIKEN), Genomic
            Sciences Center
            3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
            Tel: +81-298-36-9013
            Fax: +81-298-36-9098
            Email: genome-res@rtc.riken.go.jp,
            URL:http://genome.rtc.riken.go.jp/
            Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki
            ,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
            Thermostabilization and thermoactivation of thermolabile enzymes by
            trehalose and its application for the synthesis of full length
            cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
             Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
            Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki
            ,Y. and Hayashizaki,Y.
```

Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (http://genome.rtc.riken.go.jp) for further details.

FEATURES Location/Qualifiers
source 1. .189
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="A030005O22"
/clone_lib="RIKEN full-length enriched, 6 days neonate skin"
/tissue_type="skin"
/dev_stage="6 days neonate"
/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3']. cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. "

BASE COUNT 61 a 37 c 34 g 56 t 1 others
ORIGIN

Query Match 66.9%; Score 19.4; DB 161; Length 189;
Best Local Similarity 41.4%; Pred. No. 2.2e+02;
Matches 12; Conservative 11; Mismatches 6; Indels 0; Gaps 0;

```
Qy      1 uaugauucuuuuuguaagcccuaggggcu 29
          : :| ::|: :::|:||| ||: |||| :
Db    132 TTTGGTTCTATTTGTAAGACCTTGGGGGT 104
```

RESULT 10
AW939343
LOCUS AW939343 367 bp mRNA EST 30-MAY-2000
DEFINITION QV1-DT0071-090200-064-h06 DT0071 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW939343
VERSION AW939343.1 GI:8114789
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 367)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=QV1-DT0071-090200-064-h06&t3=2000-02-09&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 367.

FEATURES Location/Qualifiers
source 1. .367
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="DT0071"
/dev_stage="Adult"
/note="Organ: denis_drash; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 106 a 76 c 70 g 113 t 2 others
ORIGIN

Query Match 66.9%; Score 19.4; DB 122; Length 367;
Best Local Similarity 48.3%; Pred. No. 2.4e+02;
Matches 14; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

```
Qy      1 uaugauucuuuuuguaagcccuaggggcu 29
          :| | ::|:: ::| |||| |:|||| |:
Db    128 TAAGTTTCTTCTTGAAAGCTCTAGGGTCT 156
```

RESULT 11
AZ579747
LOCUS AZ579747 666 bp DNA GSS 13-DEC-2000
DEFINITION 1M0367H06R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0367H06 R, DNA sequence.
ACCESSION AZ579747
VERSION AZ579747.1 GI:11694176
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 666)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0367 row: H column: 06
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 666.

FEATURES Location/Qualifiers
source 1. .666
/organism="Mus musculus"
/strain="C57BL/6J"

```
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0367H06"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT      145 a     163 c     119 g     239 t
ORIGIN


  Query Match             66.9%;  Score 19.4;  DB 246;  Length 666;
  Best Local Similarity   48.3%;  Pred. No. 2.6e+02;
  Matches   14;  Conservative    9;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 uaugauucuuuuuguaagcccuaggggcu 29
          :|  |::|::: :|   ||||:||||||:
Db     52 TACCATTCTTTCTGCTGGCCCTAGGGGCT 80


RESULT  12
AQ239971/c
LOCUS       AQ239971      309 bp    DNA             GSS       30-SEP-1998
DEFINITION  CIT-HSP-2386F6.TR.1 CIT-HSP Homo sapiens genomic clone 2386F6, DNA
            sequence.
ACCESSION   AQ239971
VERSION     AQ239971.1  GI:3672169
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 309)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
            Venter,J.C.
  TITLE     Use of a random human BAC End Sequence Database for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CIT-HSP-2386F6.TF.1
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .309
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="2386F6"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT      105 a     56 c     58 g     90 t
ORIGIN


  Query Match             66.2%;  Score 19.2;  DB 226;  Length 309;
  Best Local Similarity   54.2%;  Pred. No. 2.9e+02;
  Matches   13;  Conservative    8;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      6 uucuuuuuguaagcccuaggggcu 29
          ::|: :::|:|||||| | ||||:
Db     80 TTCTATTTGTAAGCCCCATGGGCT 57


RESULT  13
AQ072238/c
LOCUS       AQ072238      348 bp    DNA             GSS       05-AUG-1998
DEFINITION  HS_3024_B1_F02_T7 CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=3024 Col=3 Row=L, DNA sequence.
ACCESSION   AQ072238
VERSION     AQ072238.1  GI:3391087
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 348)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 3024  row: L  column: 3
            Class: BAC ends
            High quality sequence stop: 348.
FEATURES             Location/Qualifiers
     source          1. .348
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=3024 Col=3 Row=L"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /sex="male"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
BASE COUNT      111 a     59 c     49 g     129 t
ORIGIN


  Query Match             65.5%;  Score 19;  DB 224;  Length 348;
  Best Local Similarity   44.4%;  Pred. No. 3.6e+02;
  Matches   12;  Conservative   10;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      3 ugauucuuuuuguaagcccuaggggcu 29
          : |::|:::::| ||| ||:||  ||:
Db     70 TTATTCTTTTTGGAAGTCCTAGATGCT 44


RESULT  14
```

```
AI808918
LOCUS       AI808918      358 bp    mRNA            EST       07-JUL-1999
DEFINITION  wf66h01.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
            IMAGE:2360593 3', mRNA sequence.
ACCESSION   AI808918
VERSION     AI808918.1  GI:5395484
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 358)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco
            High quality sequence stop: 345.
FEATURES             Location/Qualifiers
     source          1. .358
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2360593"
                     /clone_lib="Soares_NFL_T_GBC_S1"
                     /lab_host="DH10B"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Equal amounts of plasmid DNA from three normalized
                     libraries (fetal lung NbHL19W, testis NHT, and B-cell
                     NCI_CGAP_GCB1) were mixed, and ss circles were made in
                     vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR-amplified cDNAs from pools of 5,000 clones made
                     from the same 3 libraries. The pools consisted of
                     I.M.A.G.E. clones 297480-302087, 682632-687239,
                     726408-728711, and 729096-731399. Subtraction by Bento
                     Soares and M. Fatima Bonaldo. "
BASE COUNT       86 a     75 c     69 g    127 t      1 others
ORIGIN
```

```
  Query Match            65.5%;  Score 19;  DB 102;  Length 358;
  Best Local Similarity  48.1%;  Pred. No. 3.7e+02;
  Matches   13;  Conservative    9;  Mismatches    5;  Indels    0;  Gaps    0;

        1 uaugauucuuuuuguaagcccuagggg 27
          :|:||::|: :::| || |||: || |
Db     36 TATGATTCTCTTTGGAATCCCTGGGCG 62
```

```
RESULT  15
AI047736/c
LOCUS       AI047736      408 bp    mRNA            EST       08-JUL-1998
DEFINITION  uh82f05.r1 Soares mouse urogenital ridge NMUR Mus musculus cDNA
            clone IMAGE:1764225 5', mRNA sequence.
ACCESSION   AI047736
VERSION     AI047736.1  GI:3296023
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 408)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:963749
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 386.
FEATURES             Location/Qualifiers
     source          1. .408
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1764225"
                     /clone_lib="Soares mouse urogenital ridge NMUR"
                     /sex="equal ratio of male:female"
                     /tissue_type="urogenital ridge (embryonic)"
                     /dev_stage="fetal, mixture of 11.5 and 12.5 dpc"
                     /lab_host="DH10B"
                     /note="Organ: gonad; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATTTCTTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI
                     adaptors (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     Library went through two rounds of normalization, and was
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT      117 a    102 c    102 g     87 t
ORIGIN
```

```
  Query Match            65.5%;  Score 19;  DB 15;  Length 408;
  Best Local Similarity  44.4%;  Pred. No. 3.7e+02;
  Matches   12;  Conservative   10;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      3 ugauucuuuuuguaagcccuaggggcu 29
          .  :|| :|:::::: :| | ||:||| ||:
Db    342 TGAGTCTTTTTCTAGGACCTAGGCGCT 316
```

```
Search completed: October  2, 2001, 04:38:42
Job time: 4064 sec
```

GenCore version 4.5


OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 05:03:56 ; Search time 122.48 Seconds
(without alignments)
148.670 Million cell updates/sec

Title: US-09-310-844B-24
Perfect score: 29
Sequence: 1 uaugauucuuuuuguaagcccuaggggcu 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

[redacted]tal number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_0601:*
1: /SIDS1/gcgdata/geneseq/geneseqn/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 29 | 100.0 | 29 | 21 | AAA70828 | Molecular interact |
| 2 | 29 | 100.0 | 42 | 21 | AAA71123 | Molecular interact |
| 3 | 29 | 100.0 | 42 | 21 | AAA71131 | Molecular interact |
| 4 | 29 | 100.0 | 209 | 18 | AAT61453 | IL-2 cDNA 3'end. |
| 5 | 29 | 100.0 | 209 | 21 | AAZ59456 | Interleukin-2 frag |
| 6 | 29 | 100.0 | 698 | 6 | AAN50132 | DNA sequence encod |
| 7 | 29 | 100.0 | 698 | 7 | AAN60101 | Sequence of human |
| 8 | 29 | 100.0 | 722 | 8 | AAN70144 | Sequence encoding |
| 9 | 29 | 100.0 | 769 | 11 | AAQ01764 | Human interleukin- |
| 10 | 29 | 100.0 | 784 | 5 | AAN40046 | Sequence encoding |
| 11 | 29 | 100.0 | 784 | 7 | AAN60787 | Cloned sequence en |
| 12 | 29 | 100.0 | 784 | 7 | AAN60102 | Sequence of human |
| 13 | 29 | 100.0 | 788 | 10 | AAN90253 | Interleukin-2. Ho |
| 14 | 29 | 100.0 | 794 | 5 | AAN40441 | Gene encoding poly |
| 15 | 29 | 100.0 | 794 | 10 | AAN90381 | Recombinant human |
| 16 | 29 | 100.0 | 800 | 6 | AAN50219 | Sequence encoding |
| 17 | 29 | 100.0 | 801 | 4 | AAN30031 | Sequence of interl |
| 18 | 29 | 100.0 | 801 | 5 | AAN40057 | Cloned human inter |
| 19 | 29 | 100.0 | 801 | 5 | AAN40254 | Sequence encoding |
| 20 | 29 | 100.0 | 802 | 6 | AAN50279 | DNA sequence conta |
| 21 | 29 | 100.0 | 810 | 7 | AAN60840 | Sequence encoding |
| 22 | 29 | 100.0 | 844 | 21 | AAF20965 | Human low adenosin |
| 23 | 29 | 100.0 | 844 | 21 | AAA34843 | Human adenosine re |
| 24 | 29 | 100.0 | 5561 | 21 | AAF20964 | Human low adenosin |
| 25 | 29 | 100.0 | 5561 | 21 | AAA34842 | Human adenosine re |
| 26 | 29 | 100.0 | 22421 | 21 | AAF20966 | Human low adenosin |
| 27 | 29 | 100.0 | 22421 | 21 | AAA34844 | Human adenosine re |
| 28 | 28 | 96.6 | 45 | 21 | AAA70824 | Molecular interact |
| 29 | 28 | 96.6 | 46 | 21 | AAA71087 | Molecular interact |
| 30 | 28 | 96.6 | 46 | 21 | AAA71096 | Molecular interact |
| 31 | 28 | 96.6 | 46 | 21 | AAA71099 | Molecular interact |
| 32 | 28 | 96.6 | 46 | 21 | AAA71100 | Molecular interact |
| 33 | 28 | 96.6 | 46 | 21 | AAA71104 | Molecular interact |
| 34 | 25.8 | 89.0 | 42 | 21 | AAA71113 | Molecular interact |
| 35 | 25.8 | 89.0 | 42 | 21 | AAA71118 | Molecular interact |
| 36 | 25.8 | 89.0 | 42 | 21 | AAA71126 | Molecular interact |
| 37 | 24.8 | 85.5 | 46 | 21 | AAA71085 | Molecular interact |
| 38 | 24.8 | 85.5 | 46 | 21 | AAA71103 | Molecular interact |
| 39 | 23.8 | 82.1 | 42 | 21 | AAA71114 | Molecular interact |
| 40 | 23.8 | 82.1 | 42 | 21 | AAA71119 | Molecular interact |
| 41 | 23.8 | 82.1 | 42 | 21 | AAA71127 | Molecular interact |
| 42 | 23.8 | 82.1 | 46 | 21 | AAA71094 | Molecular interact |
| 43 | 23.8 | 82.1 | 46 | 21 | AAA71110 | Molecular interact |
| 44 | 23.2 | 80.0 | 29 | 21 | AAA70829 | Molecular interact |
| 45 | 23.2 | 80.0 | 29 | 21 | AAA70830 | Molecular interact |

ALIGNMENTS

RESULT 1
AAA70828
ID AAA70828 standard; RNA; 29 BP.
XX
AC AAA70828;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #28.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Homo sapiens.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -
XX

PS Claim 235; Page 235; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 29 BP; 5 A; 5 C; 7 G; 12 U; 0 other;

Query Match 100.0%; Score 29; DB 21; Length 29;
Best Local Similarity 100.0%; Pred. No. 0.00034;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 uaugauucuuuuuguaagcccuaggggcu 29
||||||||||||||||||||||||||||||
Db 1 uaugauucuuuuuguaagcccuaggggcu 29

RESULT 2
AAA71123
ID AAA71123 standard; DNA; 42 BP.
XX
AC AAA71123;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #129.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -
XX

PS Example 7; Figure 125; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 42 BP; 9 A; 6 C; 9 G; 18 T; 0 other;

Query Match 100.0%; Score 29; DB 21; Length 42;
Best Local Similarity 58.6%; Pred. No. 0.00035;
Matches 17; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

Qy 1 uaugauucuuuuuguaagcccuaggggcu 29
:|:||::|:::::|:||||||:||||||:
Db 4 tatgattcttttttgtaagccctaggggct 32

RESULT 3
AAA71131
ID AAA71131 standard; RNA; 42 BP.
XX
AC AAA71131;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #200.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -
XX

PS Example 7; Figure 126; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 42 BP; 9 A; 6 C; 9 G; 18 U; 0 other;

Query Match 100.0%; Score 29; DB 21; Length 42;
Best Local Similarity 100.0%; Pred. No. 0.00035;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 uaugauucuuuuuguaagcccuaggggcu 29
||||||||||||||||||||||||||||||
Db 4 uaugauucuuuuuguaagcccuaggggcu 32

RESULT 4
AAT61453
ID AAT61453 standard; cDNA; 209 BP.
XX
AC AAT61453;
XX
DT 30-OCT-1997 (first entry)
XX
DE IL-2 cDNA 3'end.
XX
KW Reverse transcription; restriction enzyme; adapter; PCR;
KW amplification; isolation; T-cell; activation; expression; ss.
XX
OS Synthetic.
XX
PN WO9705286-A1.
XX
PD 13-FEB-1997.
XX
PF 30-JUL-1996; 96WO-US12468.
XX
PR 01-AUG-1995; 95US-0510032.
XX
PA (UYYA ) UNIV YALE.
XX
PI Prashar Y, Weissman S;
XX
DR WPI; 1997-145720/13.
XX
PT Isolating DNA complementary to mRNA in a sample - by reverse
PT transcription, restriction enzyme cutting, ligation to a Y-shaped
PT adapter and selective PCR amplification
XX
PS Example 1; Page 13; 46pp; English.
XX
CC The sequences given in AAT61442 to AAT61453 are claimed and are used
CC to exemplify the new method for isolating DNA in a nucleic acid
CC sample. RNA expression during early T-cell activation was studied
CC as test system.
XX
SQ Sequence 209 BP; 56 A; 30 C; 33 G; 90 T; 0 other;

Query Match 100.0%; Score 29; DB 18; Length 209;
Best Local Similarity 58.6%; Pred. No. 0.00044;
Matches 17; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

Qy 1 uaugauucuuuuuguaagcccuaggggcu 29
:|:||::|:::::|:||||||:||||||:
Db 33 tatgattctttttgtaagccctaggggct 61

RESULT 5
AAZ59456
ID AAZ59456 standard; DNA; 209 BP.
XX
AC AAZ59456;
XX
DT 11-APR-2000 (first entry)
XX
DE Interleukin-2 fragment nucleotide sequence.
XX
KW Selective amplification method; gene expression; analysis; interleukin-2;
KW mRNA level comparison; IL-2; human; ss.
XX
OS Homo sapiens.
XX
PN US6010850-A.
XX
PD 04-JAN-2000.
XX
PF 30-JUL-1996; 96US-0688514.
XX
PR 01-AUG-1995; 95US-0510032.
XX
PA (UYYA ) UNIV YALE.
XX
PI Prashar Y, Weissman SM;
XX
DR WPI; 2000-105493/09.
XX
PT DNA amplification method useful for studying changes in gene expression
PT between cell populations -
XX
PS Example 1; Column 7-8; 23pp; English.
XX
CC This is a fragment of the interleukin-2 nucleotide sequence, which can be
CC amplified by the method of the invention. The invention relates to a
CC method for selectively amplifying DNA, where the DNA has a sequence
CC complementary to a 3' end of an mRNA comprising ligating digested cDNA to
CC a adaptor. The method mediates selective PCR amplification of cDNAs under
CC high stringency. The method can be used for comparing the levels of mRNA
CC expression in two cell populations by selectively amplifying a DNA
CC fragment in a nucleic acid sample from each cell and comparing the
CC amounts of amplified fragments obtained. The comparison can be carried
CC out where at least one of the primers is labelled, preferably with a
CC radiolabel or fluorescent label, and where one of the cell populations is
CC a treated population (the other being a control population).
XX
SQ Sequence 209 BP; 56 A; 30 C; 33 G; 90 T; 0 other;

Query Match 100.0%; Score 29; DB 21; Length 209;
Best Local Similarity 58.6%; Pred. No. 0.00044;
Matches 17; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

Qy 1 uaugauucuuuuuguaagcccuaggggcu 29

```
          :|:||::|:::::|:||||||:||||||:
Db      33 tatgattctttttgtaagccctaggggct 61


RESULT   6
AAN50132
ID   AAN50132 standard; DNA; 698 BP.
XX
AC   AAN50132;
XX
DT   15-OCT-1991  (first entry)
XX
DE   DNA sequence encoding interleukin-2.
XX
KW   Interleukin-2; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             10..391
FT                   /*tag=  a
FT                   /label= interleukin-2
XX
PN   EP152358-A.
XX
PD   21-AUG-1985.
XX
PF   16-FEB-1984;   84EP-0002350.
XX
PR   16-FEB-1984;   84FR-0002350.
XX
PA   (TRAN-) TRANSGENE SA.
XX
PI   Lemoine Y,  Sondermeyer P,  Loison G,  Aigle M,  Lecocq JP;
XX
DR   WPI; 1985-204923/34.
DR   P-PSDB; AAP50111.
XX
PT   Vectors for expression of interleukin-2 yeasts - contg. yeast
PT   promoter and sequence coding for interleukin.
XX
PS   Disclosure; Fig 2; 18pp; French.
XX
CC   The human interleukin-2 gene is cloned and expressed in
CC   Saccharomyces cerevisiae in a mature form without post-translational
CC   modification.
XX
SQ   Sequence 698 BP; 251 A; 119 C; 106 G; 222 T; 0 other;


  Query Match             100.0%;  Score 29;  DB 6;  Length 698;
  Best Local Similarity   58.6%;  Pred. No. 0.00052;
  Matches   17;  Conservative   12;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 uaugauucuuuuuguaagcccuaggggcu 29
           :|:||::|:::::|:||||||:||||||:
Db     534 tatgattctttttgtaagccctaggggct 562


RESULT   7
AAN60101
ID   AAN60101 standard; cDNA; 698 BP.
XX
AC   AAN60101;
XX
DT   02-JUL-1991  (first entry)
XX
DE   Sequence of human interleukin 2 (IL2) gene on plasmid pTG 26.
XX
KW   Pox virus vector; vaccinia; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             10..393
FT                   /*tag=  a
FT   conflict        (T_45,G_45)
FT                   /*tag=  b
FT                   /citation= 1
FT   conflict        (G_387,T_387)
FT                   /*tag=  c
FT                   /citation= 1
XX
PN   EP206939-A.
XX
PD   30-DEC-1986.
XX
PF   19-JUN-1986;   86EP-0401349.
XX
PR   21-JUN-1985;   85FR-0009480.
XX
PA   (TRAN-) TRANSGENE SA.
PA   (KIEN/) KIENY M-P.
XX
PI   Kieny MP,  Sondermeyer P,  Lecocq JP;
XX
DR   WPI; 1986-341222/52.
DR   P-PSDB; AAP60088.
XX
PT   Genetically modified pox viruses - contg. DNA coding for human
PT   interleukin 2
XX
PS   Example ; Fig 1; 33pp; French.
XX
CC   cDNA coding for IL2 in plasmid pTG26 is restructured by inserting a
CC   92bp synthetic peptide signal fragment, giving a plasmid pTG36. A
CC   plasmid pTG186-POLY is prepd. as in EP 206920. This is digested with
CC   PstI and ligated with PstI-digested pTG36. The recombinants are used
CC   to transform E.coli and selected to give plasmid pVVIL2 (pTG188)
CC   which is cloned in vaccinia virus. Citation 1 = Taniguchi,T. et al.
CC   (1983) Nature,302,p305. The conflicts are silent wrt translation.
XX
SQ   Sequence 698 BP; 251 A; 117 C; 108 G; 222 T; 0 other;


  Query Match             100.0%;  Score 29;  DB 7;  Length 698;
  Best Local Similarity   58.6%;  Pred. No. 0.00052;
  Matches   17;  Conservative   12;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 uaugauucuuuuuguaagcccuaggggcu 29
           :|:||::|:::::|:||||||:||||||:
Db     534 tatgattctttttgtaagccctaggggct 562


RESULT   8
AAN70144
ID   AAN70144 standard; cDNA; 722 BP.
XX
AC   AAN70144;
XX
DT   17-APR-1991  (first entry)
XX
DE   Sequence encoding human interleukin-2 (IL-2).
XX
KW   Mastitis therapy; parasitic infection; lymphokine; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             8..67
FT                   /*tag=  a
FT   mat_peptide     68..469
FT                   /*tag=  b
FT   misc_feature    11..719
FT                   /*tag=  c
FT                   /note= "used to probe bovine cDNA library"
```

XX
PN EP215576-A.
XX
PD 25-MAR-1987.
XX
PF 15-AUG-1986; 86EP-0306303.
XX
PR 31-JUL-1986; 86US-0888994.
PR 16-AUG-1985; 85US-0766643.
XX
PA (IMMU-) IMMUNEX CORP.
XX
PI Anderson DM, Baker PE, Cantrell MA, Cerretti DP, Cosman DJ;
PI Gimpel SD, Grabstein KH, Larsen AD, McKereghan KN;
XX
DR WPI; 1987-081523/12.
DR P-PSDB; AAP70090.
XX
PT Recombinant bovine interleukin-2 - for treating mastitis,
PT respiratory and gastro-intestinal syndromes and parasitic
PT infections
XX
PS Disclosure; Fig 1; 47pp; English.
XX
CC A bovine cDNA library was screened using a probe which was essentially
CC the entire length of human IL-2 cDNA (AAN70144). A single positive
CC host colony was identified. Plasmid bIL-2-4 was prepd. and the
CC nuclectide sequence detd. (AAN70145). The specific activity of the
CC bIL-2 is 4.5 x 10(4) units/ug protein.
XX
SQ Sequence 722 BP; 245 A; 130 C; 112 G; 235 T; 0 other;

Query Match 100.0%; Score 29; DB 8; Length 722;
Best Local Similarity 58.6%; Pred. No. 0.00052;
Matches 17; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

Qy 1 uaugauucuuuuuguaagcccuaggggcu 29
:|:||::|:::::|:|||||||:|||||:
Db 610 tatgattctttttgtaagccctaggggct 638

RESULT 9
AAQ01764
ID AAQ01764 standard; DNA; 769 BP.
XX
AC AAQ01764;
XX
DT 27-JUL-1990 (first entry)
XX
DE Human interleukin-2 gene.
XX
KW IL-2; dhfr; dihydrofolic acid reductase; differentiation.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 3..506
FT /*tag= a
XX
PN JP02009388-A.
XX
PD 12-JAN-1990.
XX
PF 08-JUL-1988; 88JP-0170142.
XX
PR 09-MAR-1988; 88JP-0055270.
XX
PA (AJIN ) AJINOMOTO KK.
XX
DR WPI; 1990-055348/08.
DR P-PSDB; AAR05414.
XX
PT Physiologically active protein prepn. -
PT by transforming plasmid having gene coding physiologically
PT active protein and gene of di:hydrofolic acid reductase to hamster
PT ovary etc.
XX
PS Example 2; Fig 4; 12pp; Japanese.
XX
CC Gene may be expressed by transforming a dhfr negative strain of CHO cells
CC with an active IL-2 gene and dhfr carrying vector.
XX
SQ Sequence 769 BP; 280 A; 131 C; 113 G; 245 T; 0 other;

Query Match 100.0%; Score 29; DB 11; Length 769;
Best Local Similarity 58.6%; Pred. No. 0.00053;
Matches 17; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

Qy 1 uaugauucuuuuuguaagcccuaggggcu 29
:|:||::|:::::|:|||||||:|||||:
Db 605 tatgattctttttgtaagccctaggggct 633

RESULT 10
AAN40046
ID AAN40046 standard; cDNA; 784 BP.
XX
AC AAN40046;
XX
DT 14-JAN-1992 (first entry)
XX
DE Sequence encoding interleukin-2 (IL-2) related polypeptide in
DE clone pSV-hIL-2L-O.
XX
KW Diagnosis; therapy; cancer; tumour-specific cytotoxic cell; AIDS;
KW multiple sclerosis; lupus; rheumatoid arthritis; herpes;
KW viral disease; lymphokine; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 48..509
FT /*tag= a
FT polyA_signal 771..776
FT /*tag= b
XX
PN EP118977-A.
XX
PD 19-SEP-1984.
XX
PF 25-JAN-1984; 84EP-0300439.
XX
PR 10-JUN-1983; 83GB-0015981.
PR 08-FEB-1983; 83GB-0003383.
XX
PA (BIOJ ) BIOGEN NV.
XX
PI Fiers WC, Devos RR;
XX
DR WPI; 1984-232548/38.
DR P-PSDB; AAP40055.
XX
PT Prodn. of human interleukin 2-like polypeptide(s) - useful
PT instead of IL-2 for stimulating the immune system etc.
XX
PS Disclosure; Fig 4; 69pp; English.
XX
CC The DNA sequence is esp. selected from a human chromosomal gene bank,
CC e.g. it is a hIL-2 related portion of lamda CH4A-ghIL-2-1 or -2, or
CC of lamda L47-ghIL-2-1, -2 or -3. Transformed hosts are also claimed,
CC esp. E.coli, Ps spp., B.subtilis, B.stearothermophilus. IL-2-like
CC polypeptides are also claimed. AAN40046 is an IL-2 related DNA SQ of
CC the invention. It is missing at least part of the coding region for
CC the amino acids of the putaive signal sequence of pre hIL-2 and the

```
CC   ATG start signal of that signal sequence.
XX
SQ   Sequence 784 BP; 267 A; 146 C; 115 G; 256 T; 0 other;


  Query Match            100.0%;  Score 29;  DB 5;  Length 784;
  Best Local Similarity   58.6%;  Pred. No. 0.00053;
  Matches   17;  Conservative   12;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 uaugauucuuuuuguaagcccuaggggcu 29
          :|:||::|:::::|:|||||:||||||:
Db    650 tatgattctttttgtaagccctaggggct 678


RESULT  11
AAN60787
ID   AAN60787 standard; cDNA; 784 BP.
XX
AC   AAN60787;
XX
DT   03-JUL-1991  (first entry)
XX
DE   Cloned sequence encoding human T-cell growth factor.
XX
KW   aTCGF; antitumour; cancer; IL-2; immunoregulatory; AIDS;
KW   cytomegalovirus; herpes; tuberculosis; leptosis; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             43..504
FT                   /*tag=  a
FT   sig_peptide     43..101
FT                   /*tag=  b
FT   mat_peptide     102..501
FT                   /*tag=  c
XX
PN   WO8505124-A.
XX
PD   21-NOV-1985.
XX
PF   02-MAY-1985;   85WO-EP00194.
XX
PR   08-MAY-1984;   84US-0608228.
PR   07-APR-1986;   86US-0849234.
XX
PA   (SANO ) SANDOZ AG.
PA   (GENE-) GENETICS INST INC.
XX
AU   Clark SC;
XX
DR   WPI; 1986-028179/04.
DR   P-PSDB; AAP60902.
XX
PT   New human T-cell growth factor from m-RNA of plasma source -
PT   useful as immuno-regulatory agent for therapeutic use, as
PT   antimycotic and antiviral agent, for assisting antitumour
PT   treatments etc.
XX
PS   Example A; Page 27; 66pp; English.
XX
CC   Human T-cell growth factor (aTCGF) is distinct from TCGF (IL-2) may
CC   be purified and recovered with a higher yeild. It is useful in
CC   supporting activated mammalian T-cell growth eg. in culture media.
CC   It may also be used as an antimycotic, immunoregulatory antiviral
CC   and antibacterial agent treating AIDS, SCIDS, congenital immuno-
CC   deficiency, cytomegalovirus and herpes infections, tuberculosis and
CC   leptosis.
XX
SQ   Sequence 784 BP; 273 A; 143 C; 114 G; 254 T; 0 other;


  Query Match            100.0%;  Score 29;  DB 7;  Length 784;
  Best Local Similarity   58.6%;  Pred. No. 0.00053;
  Matches   17;  Conservative   12;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 uaugauucuuuuuguaagcccuaggggcu 29
          :|:||::|:::::|:|||||:||||||:
Db    641 tatgattctttttgtaagccctaggggct 669


RESULT  12
AAN60102
ID   AAN60102 standard; DNA; 784 BP.
XX
AC   AAN60102;
XX
DT   02-JUL-1991  (first entry)
XX
DE   Sequence of human interleukin 2 (IL2) gene and signal fragment
DE   on pTG36.
XX
KW   Pox virus vector; vaccinia; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             13..474
FT                   /*tag=  a
XX
PN   EP206939-A.
XX
PD   30-DEC-1986.
XX
PF   19-JUN-1986;   86EP-0401349.
XX
PR   21-JUN-1985;   85FR-0009480.
XX
PA   (TRAN-) TRANSGENE SA.
PA   (KIEN/) KIENY M-P.
XX
PI   Kieny MP,  Sondermeyer P,  Lecocq JP;
XX
DR   WPI; 1986-341222/52.
DR   P-PSDB; AAP60089.
XX
PT   Genetically modified pox viruses - contg. DNA coding for human
PT   interleukin 2
XX
PS   Example ; Fig 5; 33pp; French.
XX
CC   cDNA coding for IL2 in plasmid pTG26 is restructured by inserting a
CC   92bp synthetic peptide signal fragment, giving a plasmid pTG36. A
CC   plasmid pTG186-POLY is prepd. as in EP 206920. This is digested with
CC   PstI and ligated with PstI-digested pTG36. The recombinants are used
CC   to transform E.coli and selected to give plasmid pVVIL2 (pTG188)
CC   which is cloned in vaccinia virus.
XX
SQ   Sequence 784 BP; 273 A; 148 C; 116 G; 247 T; 0 other;


  Query Match            100.0%;  Score 29;  DB 7;  Length 784;
  Best Local Similarity   58.6%;  Pred. No. 0.00053;
  Matches   17;  Conservative   12;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 uaugauucuuuuuguaagcccuaggggcu 29
          :|:||::|:::::|:|||||:||||||:
Db    615 tatgattctttttgtaagccctaggggct 643


RESULT  13
AAN90253
ID   AAN90253 standard; DNA; 788 BP.
XX
AC   AAN90253;
XX
```

```
DT   01-NOV-1989  (first entry)
XX
DE   Interleukin-2.
XX
KW   Interleukin-2; lysine-depleted variant; site-directed
KW   mutagenesis; human.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             43..501
FT                   /*tag=  a
XX
PN   WO8905824-A.
XX
PD   29-JUN-1989.
XX
PF   22-DEC-1988;   88WO-US04633.
XX
PR   23-DEC-1987;   87US-0137043.
XX
PA   (GENE ) GENETICS INST INC.
XX
PI   Shaw G;
XX
DR   WPI; 1989-206594/28.
DR   P-PSDB; AAP90467.
XX
PT   New lysine depleted variants of polypeptide
PT   - opt. modified with hydrophilic residues,
PT   biologically active but with altered solubility, stability etc.
XX
PS   Disclosure; fig 1; 35pp; English.
XX
CC   DNA of interleukin-2 (see corresp. AAP90467). Used in the patent
CC   to create lysine depleted variants by site-directed mutagenesis, or
CC   synthesis.
XX
SQ   Sequence 788 BP; 272 A; 146 C; 113 G; 257 T; 0 other;


  Query Match             100.0%;  Score 29;  DB 10;  Length 788;
  Best Local Similarity   58.6%;  Pred. No. 0.00053;
  Matches   17;  Conservative   12;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 uaugauucuuuuuguaagcccuaggggcu 29
           :|:||::|:::::|:|||||||:|||||:
Db     645 tatgattctttttgtaagccctaggggct 673


RESULT  14
AAN40441
ID   AAN40441 standard; cDNA; 794 BP.
XX
AC   AAN40441;
XX
DT   21-JUL-1992  (first entry)
XX
DE   Gene encoding polypeptide having IL-2 activity.
XX
KW   Interleukin.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             47..508
FT                   /*tag=  a
FT   PolyA_signal    770..775
FT                   /*tag=  b
FT   PolyA_site      788..794
FT                   /*tag=  c
XX
PN   JP59139349-A.
XX
PD   10-AUG-1984.
XX
PF   29-DEC-1982;   82JP-0230372.
XX
PR   29-DEC-1982;   82JP-0230372.
XX
PA   (AJIN ) AJINOMOTO KK.
PA   (GANK-) GAN KENKYUKAI ZH.
XX
DR   WPI; 1984-233967/38.
XX
PT   Polypeptide having interleukin-2 activity - with amino acid sequence
PT   having C-terminal threonine.
XX
PS   Disclosure; Page 7; 16pp; Japanese.
XX
CC   The DNA was obtd. from mRNA of human leukaemia cells.
XX
SQ   Sequence 794 BP; 276 A; 147 C; 113 G; 258 T; 0 other;


  Query Match             100.0%;  Score 29;  DB 5;  Length 794;
  Best Local Similarity   58.6%;  Pred. No. 0.00053;
  Matches   17;  Conservative   12;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 uaugauucuuuuuguaagcccuaggggcu 29
           :|:||::|:::::|:|||||||:|||||:
Db     649 tatgattctttttgtaagccctaggggct 677


RESULT  15
AAN90381
ID   AAN90381 standard; cDNA; 794 BP.
XX
AC   AAN90381;
XX
DT   03-NOV-1989  (first entry)
XX
DE   Recombinant human interleukin-2 DNA.
XX
KW   Recombinant human interleukin-2 DNA; jurkat cells; leukaemia.
XX
OS   Homo sapiens (human).
XX
FH   Key             Location/Qualifiers
FT   sig_peptide     47..106
FT                   /*tag=  a
FT   mat_peptide     107..505
FT                   /*tag=  b
XX
PN   JP01165399-A.
XX
PD   29-JUN-1989.
XX
PF   24-DEC-1982;   82JP-0292084.
XX
PR   24-DEC-1982;   82JP-0292084.
XX
PA   (AJIN ) AJINOMOTO KK.
PA   (GANK-) ZH GAN KENKYUKAI.
XX
DR   WPI; 1989-230772/32.
DR   P-PSDB; AAP90614.
XX
PT   Recombinant human interleukin-2
PT   - produced by irradiating hIL-2-producing
PT   jurkat cells, extracting mRNA, forming cDNA, inserting into
PT   pBR322 and transforming E. coli.
XX
PS   Claim 1; page 737; 16pp; Japanese.
XX
CC   Recombinant human interleukin-2 (hIL-2) DNA (see AAP90614). Isolated
```

CC from jurkat cells of human leukaemic T lymphocytes. The jurkat
CC cells are treated with X-rays and cultured in the presence of
CC concanavalin A to obtain variants which produce 40 times more IL
CC than the original jurkat cells.
XX
SQ Sequence 794 BP; 276 A; 146 C; 114 G; 258 T; 0 other;

Query Match 100.0%; Score 29; DB 10; Length 794;
Best Local Similarity 58.6%; Pred. No. 0.00053;
Matches 17; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 uaugauucuuuuuguaagcccuaggggcu 29
           :|:||::|:::::|:||||||:||||||:
Db     649 tatgattctttttgtaagccctaggggct 677
```

Search completed: October 2, 2001, 05:03:57
Job time: 3809 sec

GenCore version 4.5


OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 04:38:42 ; Search time 1643.83 Seconds
(without alignments)
166.765 Million cell updates/sec

Title: US-09-310-844B-25
Perfect score: 29
Sequence: 1 aaagauucuuuuuguaagccccaagggcu 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est33:*
26: gb_est34:*
27: gb_est35:*
28: gb_est36:*
29: gb_est37:*
30: gb_est38:*
31: gb_est39:*
32: gb_est40:*
33: em_estba:*
34: em_estfun:*
35: em_esthum1:*
36: em_esthum2:*
37: em_esthum3:*
38: em_esthum4:*
39: em_esthum5:*
40: em_esthum6:*
41: em_esthum7:*
42: em_esthum8:*
43: em_esthum9:*
44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estin5:*
68: em_estom1:*
69: em_estom2:*
70: em_estov1:*
71: em_estov2:*
72: em_estpl1:*
73: em_estpl2:*
74: em_estpl3:*
75: em_estpl4:*
76: em_estpl5:*
77: em_estpl6:*
78: em_estpl7:*
79: em_estpl8:*
80: em_estpl9:*
81: em_estpl10:*
82: em_estro1:*
83: em_estro2:*
84: em_estro3:*
85: em_estro4:*
86: em_estro5:*
87: em_estro6:*
88: em_estro7:*
89: em_estro8:*
90: em_estro9:*
91: em_estro10:*
92: em_estro11:*
93: em_estro12:*
94: em_estro13:*
95: em_estro14:*
96: em_estro15:*
97: em_estro16:*
98: em_estro17:*
99: em_estro18:*
100: em_estro19:*
101: em_estro20:*
102: gb_est25:*
103: gb_est26:*
104: gb_est27:*
105: gb_est28:*
106: gb_est29:*
107: gb_est30:*
108: gb_est31:*
109: gb_est32:*
110: gb_est41:*
111: gb_est42:*
112: gb_est43:*
113: gb_est44:*
114: gb_est45:*
115: gb_est46:*
116: gb_est47:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est75:*
137: gb_est76:*
138: gb_est77:*
139: gb_est78:*
140: gb_est79:*
141: gb_est80:*
142: gb_est81:*
143: gb_est82:*
144: gb_est83:*
145: gb_est84:*
146: gb_est85:*
147: gb_est86:*
148: gb_est87:*
149: gb_est88:*
150: gb_est89:*
151: gb_est90:*
152: gb_est99:*
153: gb_est100:*
154: gb_est101:*
155: gb_est102:*
156: gb_est103:*
157: gb_est104:*
158: gb_est105:*
159: gb_est106:*
160: gb_est67:*
161: gb_est68:*
162: gb_est69:*
163: gb_est70:*
164: gb_est71:*
165: gb_est72:*
166: gb_est73:*
167: gb_est74:*
168: gb_est91:*
169: gb_est92:*
170: gb_est93:*
171: gb_est94:*
172: gb_est95:*
173: gb_est96:*
174: gb_est97:*
175: gb_est98:*
176: em_esthum29:*
177: em_esthum30:*
178: em_esthum31:*
179: em_esthum32:*
180: em_esthum33:*
181: em_estom3:*
182: em_estpl11:*
183: em_estro21:*
184: em_estro22:*
185: em_estro23:*
186: em_htc:*
187: gb_est107:*
188: gb_est108:*
189: gb_est109:*
190: gb_est110:*
191: gb_est111:*
192: gb_htc:*
193: em_gss_fun:*
194: em_gss_hum1:*
195: em_gss_hum2:*
196: em_gss_hum3:*
197: em_gss_hum4:*
198: em_gss_hum5:*
199: em_gss_hum6:*
200: em_gss_hum7:*
201: em_gss_hum8:*
202: em_gss_hum9:*
203: em_gss_inv1:*
204: em_gss_inv2:*
205: em_gss_inv3:*
206: em_gss_other:*
207: em_gss_pln1:*
208: em_gss_pln2:*
209: em_gss_pro:*
210: em_gss_rod1:*
211: em_gss_rod2:*
212: em_gss_rod3:*
213: em_gss_rod4:*
214: em_gss_rod5:*
215: em_gss_vrt1:*
216: em_gss_vrt2:*
217: em_gss_vrt3:*
218: gb_gss1:*
219: gb_gss2:*
220: gb_gss3:*
221: gb_gss4:*
222: gb_gss5:*
223: gb_gss6:*
224: gb_gss7:*
225: gb_gss8:*
226: gb_gss9:*
227: gb_gss10:*
228: gb_gss11:*
229: gb_gss12:*
230: gb_gss13:*
231: gb_gss14:*
232: gb_gss15:*
233: gb_gss16:*
234: gb_gss17:*
235: gb_gss18:*
236: gb_gss19:*
237: gb_gss20:*
238: gb_gss21:*
239: gb_gss22:*
240: gb_gss23:*
241: gb_gss24:*
242: gb_gss25:*
243: gb_gss26:*
244: gb_gss27:*
245: gb_gss28:*
246: gb_gss29:*
247: gb_gss30:*
248: gb_gss31:*
249: gb_gss32:*
250: gb_gss33:*
251: gb_gss34:*
252: em_gss_inv4:*
253: em_gss_rod6:*
254: em_gss_rod7:*
255: em_gss_rod8:*
256: gb_gss35:*
257: gb_gss36:*
258: gb_gss37:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|---|
| 1 | | 27 | 93.1 | 325 | 4 | AA290106 | AA290106 vb32e03.r |
| 2 | c | 23.2 | 80.0 | 527 | 226 | AQ267968 | AQ267968 RPCI11-73 |
| 3 | c | 23.2 | 80.0 | 596 | 226 | AQ267210 | AQ267210 RPCI11-73 |
| 4 | c | 21 | 72.4 | 246 | 152 | BG365004 | BG365004 112339 MA |
| 5 | c | 21 | 72.4 | 516 | 233 | AQ809431 | AQ809431 HS_3149_B |
| 6 | c | 20.8 | 71.7 | 309 | 226 | AQ239971 | AQ239971 CIT-HSP-2 |
| 7 | | 20 | 69.0 | 646 | 243 | AZ407430 | AZ407430 1M0178F16 |
| 8 | c | 19.8 | 68.3 | 470 | 230 | AQ592351 | AQ592351 HS_5446_A |
| 9 | | 19.8 | 68.3 | 679 | 229 | AQ487601 | AQ487601 RPCI-11-2 |
| 10 | | 19.8 | 68.3 | 694 | 229 | AQ487546 | AQ487546 RPCI-11-2 |
| 11 | | 19.8 | 68.3 | 723 | 229 | AQ487597 | AQ487597 RPCI-11-2 |
| 12 | c | 19.8 | 68.3 | 1101 | 219 | CNS00GMB | AL071964 Drosophil |
| 13 | | 19.6 | 67.6 | 332 | 119 | AW661763 | AW661763 hi80e01.x |
| 14 | | 19.6 | 67.6 | 350 | 11 | AA761822 | AA761822 nz29a05.s |
| 15 | | 19.6 | 67.6 | 428 | 230 | AQ561664 | AQ561664 HS_5206_A |
| 16 | c | 19.6 | 67.6 | 449 | 223 | AQ022080 | AQ022080 CIT-HSP-2 |
| 17 | c | 19.6 | 67.6 | 518 | 229 | AQ515781 | AQ515781 HS_5237_A |
| 18 | c | 19.6 | 67.6 | 683 | 32 | AV700573 | AV700573 AV700573 |
| 19 | c | 19.6 | 67.6 | 926 | 220 | CNS01ZC8 | AL174113 Tetraodon |
| 20 | c | 19.4 | 66.9 | 203 | 162 | BB604151 | BB604151 BB604151 |
| 21 | | 19.4 | 66.9 | 334 | 251 | AZ898059 | AZ898059 RPCI-24-1 |
| 22 | | 19.4 | 66.9 | 464 | 24 | AI786847 | AI786847 uj32d02.y |
| 23 | c | 19.4 | 66.9 | 477 | 223 | AQ001016 | AQ001016 CIT-HSP-2 |
| 24 | | 19.4 | 66.9 | 482 | 231 | AQ618671 | AQ618671 HS_5169_A |
| 25 | c | 19.4 | 66.9 | 524 | 23 | AI649602 | AI649602 486068C02 |
| 26 | | 19.4 | 66.9 | 620 | 250 | AZ856295 | AZ856295 2M0160J08 |
| 27 | | 19.4 | 66.9 | 819 | 222 | CNS056V4 | AL323833 Tetraodon |
| 28 | c | 19.4 | 66.9 | 872 | 239 | AZ184475 | AZ184475 SP_1003_A |
| 29 | | 19.4 | 66.9 | 877 | 238 | AZ128746 | AZ128746 OSJNBb009 |
| 30 | c | 19.2 | 66.2 | 107 | 244 | AZ481896 | AZ481896 1M0306G03 |
| 31 | | 19.2 | 66.2 | 393 | 21 | AI506956 | AI506956 vl39d03.x |
| 32 | | 19.2 | 66.2 | 541 | 239 | AZ161007 | AZ161007 SP_0067_B |
| 33 | | 19.2 | 66.2 | 566 | 250 | AZ818819 | AZ818819 2M0089N04 |
| 34 | c | 19.2 | 66.2 | 571 | 188 | R84287 | R84287 yp89d08.r1 |
| 35 | | 19.2 | 66.2 | 691 | 250 | AZ818818 | AZ818818 2M0089N02 |
| 36 | c | 19 | 65.5 | 348 | 23 | AI686817 | AI686817 tp79f12.x |
| 37 | | 19 | 65.5 | 424 | 103 | AI898135 | AI898135 EST267578 |
| 38 | | 19 | 65.5 | 430 | 232 | AQ674237 | AQ674237 HS_5490_A |
| 39 | | 19 | 65.5 | 443 | 3 | AA193050 | AA193050 zp94e03.r |
| 40 | | 19 | 65.5 | 452 | 225 | AQ224588 | AQ224588 HS_2012_A |
| 41 | | 19 | 65.5 | 456 | 249 | AZ797512 | AZ797512 2M0053H14 |
| 42 | c | 19 | 65.5 | 482 | 251 | AZ891696 | AZ891696 RPCI-24-1 |
| 43 | | 19 | 65.5 | 491 | 250 | AZ848492 | AZ848492 2M0149F02 |
| 44 | | 19 | 65.5 | 552 | 250 | AZ813750 | AZ813750 2M0081K07 |
| 45 | c | 19 | 65.5 | 579 | 248 | AZ714234 | AZ714234 RPCI-24-8 |

ALIGNMENTS

```
RESULT   1
AA290106
LOCUS       AA290106      325 bp    mRNA            EST       14-APR-1997
DEFINITION  vb32e03.r1 Soares mouse lymph node NbMLN Mus musculus cDNA clone
            IMAGE:750652 5' similar to gb:M16762 Mouse interleukin 2 (MOUSE);,
            mRNA sequence.
ACCESSION   AA290106
VERSION     AA290106.1  GI:1936335
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 325)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:459636
            Putative full length read
            vector to vector length is 422
            Seq primer: -28m13 rev2 ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .325
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:750652"
                     /clone_lib="Soares mouse lymph node NbMLN"
                     /sex="male"
                     /tissue_type="lymph node"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
                     /note="Organ: lymph node; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Not I - oligo(dT) primer
                     [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGATACTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. RNA
                     provided by Dr. Bertrand Jordan. Library constructed and
                     normalized by Bento Soares and M.Fatima Bonaldo."
BASE COUNT      107 a     47 c     41 g    130 t
ORIGIN


  Query Match             93.1%;  Score 27;  DB 4;  Length 325;
  Best Local Similarity   70.4%;  Pred. No. 0.13;
  Matches   19;  Conservative    8;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 aaagauucuuuuuguaagccccaaggg 27
          ||||||::|:::::|:||||||||||||
Db    167 AAAGATTCTTTTTGTAAGCCCCAAGGG 193


RESULT   2
AQ267968/c
LOCUS       AQ267968      527 bp    DNA             GSS       27-APR-1999
DEFINITION  RPCI11-73E11.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-73E11,
            DNA sequence.
ACCESSION   AQ267968
VERSION     AQ267968.1  GI:3795572
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 527)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
  TITLE     Use of human BAC End Sequences for Sequence-Ready Map Building
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: RPCI11-73E11.TK
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
```

```
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: SP6
            Class: BAC ends.
FEATURES              Location/Qualifiers
     source           1. .527
                      /organism="Homo sapiens"
                      /db_xref="GDB:7527754"
                      /db_xref="taxon:9606"
                      /clone="RPCI-11-73E11"
                      /clone_lib="RPCI-11"
                      /sex="Male"
                      /cell_type="Lymphocytes"
                      /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                      RPCI11 Human Male BAC Library"
BASE COUNT      203 a      78 c      77 g     167 t       2 others
ORIGIN


  Query Match             80.0%;  Score 23.2;  DB 226;  Length 527;
  Best Local Similarity   57.1%;  Pred. No. 6.5;
  Matches   16;  Conservative    9;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 aagauucuuuuuguaagccccaagggcu 29
           | ||::|:::::|:|||||| | ||||:
Db     357 ATGATTCTTTTTGTAAGCCCTAGGGGCT 330


RESULT   3
AQ267210/c
LOCUS       AQ267210      596 bp    DNA             GSS       27-APR-1999
DEFINITION  RPCI11-73C7.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-73C7, DNA
            sequence.
ACCESSION   AQ267210
VERSION     AQ267210.1  GI:3794814
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 596)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
  TITLE     Use of human BAC End Sequences for Sequence-Ready Map Building
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: RPCI11-73C7.TK
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: SP6
            Class: BAC ends.
FEATURES              Location/Qualifiers
     source           1. .596
                      /organism="Homo sapiens"
                      /db_xref="GDB:7527702"
                      /db_xref="taxon:9606"
                      /clone="RPCI-11-73C7"
                      /clone_lib="RPCI-11"
                      /sex="Male"
                      /cell_type="Lymphocytes"
                      /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                      RPCI11 Human Male BAC Library"
BASE COUNT      227 a      92 c      88 g     186 t       3 others
ORIGIN


  Query Match             80.0%;  Score 23.2;  DB 226;  Length 596;
  Best Local Similarity   57.1%;  Pred. No. 6.6;
  Matches   16;  Conservative    9;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 aagauucuuuuuguaagccccaagggcu 29
           | ||::|:::::|:|||||| | ||||:
Db     357 ATGATTCTTTTTGTAAGCCCTAGGGGCT 330


RESULT   4
BG365004/c
LOCUS       BG365004      246 bp    mRNA            EST       08-MAR-2001
DEFINITION  112339 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION   BG365004
VERSION     BG365004.1  GI:13254101
KEYWORDS    EST.
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 246)
  AUTHORS   Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M.,
            Bennett,G.A., Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Laegreid
            ,W.W. and Keele,J.W.
  TITLE     Design and use of four pooled tissue normalized cDNA libraries for
            EST discovery in cattle
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and alt_trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 103  row: K  column: 22
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES              Location/Qualifiers
     source           1. .246
                      /organism="Bos taurus"
                      /db_xref="taxon:9913"
                      /clone_lib="MARC 1BOV"
                      /tissue_type="pooled"
                      /lab_host="DH10B"
                      /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
                      Library made from pooled tissue from lymph node, ovary,
                      fat, hypothalamus, and pituitary."
BASE COUNT       62 a      49 c      68 g      67 t
ORIGIN


  Query Match             72.4%;  Score 21;  DB 152;  Length 246;
  Best Local Similarity   51.7%;  Pred. No. 53;
  Matches   15;  Conservative    9;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 aaagauucuuuuuguaagccccaagggcu 29
           |||| :: ::::: :|||||||||||  |:
Db     198 AAAGGTTTTTTTTTTAAGCCCCAAGCCCT 170
```

```
RESULT   5
AQ809431/c
LOCUS       AQ809431      516 bp    DNA             GSS       10-AUG-1999
DEFINITION  HS_3149_B1_A05_T7C CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=3149 Col=9 Row=B, DNA sequence.
ACCESSION   AQ809431
VERSION     AQ809431.1  GI:5728673
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 516)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones may be purchased from Research Genetics (info@resgen.com).
            BAC end Web Server: http://www.htsc.washington.edu
            Plate: 3149  row: B  column: 9
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 516.
FEATURES             Location/Qualifiers
     source          1. .516
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=3149 Col=9 Row=B"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /sex="male"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
BASE COUNT      162 a    112 c     99 g    139 t      4 others
ORIGIN

  Query Match             72.4%;  Score 21;  DB 233;  Length 516;
  Best Local Similarity   58.6%;  Pred. No. 60;
  Matches   17;  Conservative    7;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 aaagauucuuuuuguaagccccaagggcu 29
          ||| | :|: :::|:||||| |||| ||:
Db     90 AAACACTCTGTTTGTAAGCCGCAAGTGCT 62


RESULT   6
AQ239971/c
LOCUS       AQ239971      309 bp    DNA             GSS       30-SEP-1998
DEFINITION  CIT-HSP-2386F6.TR.1 CIT-HSP Homo sapiens genomic clone 2386F6, DNA
            sequence.
ACCESSION   AQ239971
VERSION     AQ239971.1  GI:3672169
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 309)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
            Venter,J.C.
  TITLE     Use of a random human BAC End Sequence Database for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CIT-HSP-2386F6.TF.1
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .309
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="2386F6"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT      105 a     56 c     58 g     90 t
ORIGIN

  Query Match             71.7%;  Score 20.8;  DB 226;  Length 309;
  Best Local Similarity   58.3%;  Pred. No. 68;
  Matches   14;  Conservative    8;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      6 uucuuuuuguaagccccaagggcu 29
          ::|: :::|:|||||||| ||||:
Db     80 TTCTATTTGTAAGCCCCATGGGCT 57


RESULT   7
AZ407430
LOCUS       AZ407430      646 bp    DNA             GSS       03-OCT-2000
DEFINITION  1M0178F16F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0178F16 F, DNA sequence.
ACCESSION   AZ407430
VERSION     AZ407430.1  GI:10531539
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 646)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0178  row: F  column: 16
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 646.
FEATURES             Location/Qualifiers
```

```
    source          1. .646
                    /organism="Mus musculus"
                    /strain="C57BL/6J"
                    /db_xref="taxon:10090"
                    /clone="UUGC1M0178F16"
                    /clone_lib="Mouse 10kb plasmid UUGC1M library"
                    /sex="Male"
                    /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                    /note="Vector: PWD42nv; Purified genomic DNA from M.
                    musculus C57BL/6J (male) was obtained  from the Jackson
                    Laboratory Mouse DNA Resource
                    (http://www.jax.org/resources/documents/dnares/). The DNA
                    was hydrodynamically sheared by repeated passage through a
                    0.005 inch orifice at constant velocity. The sheared DNA
                    was blunt end-repaired with T4 DNA polymerase and T4
                    polynucleotide kinase. Adaptor oligonucleotides were
                    ligated to the blunt ends in high  molar excess. The
                    adaptored DNA was purified and size-selected for a 9.5 to
                    10.5 kb range using preparative agarose gel
                    electrophoresis. Vector DNA was prepared from a derivative
                    of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                    inducible derivative of plasmid R1. The vector was ligated
                    with adaptors complementary to the insert adaptors and
                    purified. The sheared, adaptored mouse DNA was annealed to
                    adaptored vector DNA, and transformed into
                    chemically-competent E. coli XL10-Gold (Stratagene) cells
                    and selected for ampicillin resistance."
BASE COUNT      151 a     96 c    115 g    284 t
ORIGIN


  Query Match             69.0%;  Score 20;  DB 243;  Length 646;
  Best Local Similarity   53.6%;  Pred. No. 1.7e+02;
  Matches   15;  Conservative    8;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      2 aagauucuuuuuguaagccccaagggcu 29
          ||||::|:::::| | || |  |||||:
Db    134 AAGATTCTTTTTGGAGGCACTCAGGGCT 161


RESULT   8
AQ592351/c
LOCUS       AQ592351      470 bp    DNA             GSS       08-JUN-1999
DEFINITION  HS_5446_A1_B11_SP6E RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate=1022 Col=21 Row=C, DNA sequence.
ACCESSION   AQ592351
VERSION     AQ592351.1  GI:5024003
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 470)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
            Plate: 1022  row: C  column: 21
            Seq primer: SP6
            Class: BAC ends
            High quality sequence stop: 470.
FEATURES            Location/Qualifiers
     source         1. .470
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="Plate=1022 Col=21 Row=C"
                    /clone_lib="RPCI-11 Human Male BAC Library"
                    /sex="male"
                    /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                    Male blood DNA was isolated from one randomly chosen donor
                    and partially digested with a combination of EcoRI and
                    EcoRI Methylase. Size selected DNA was cloned into the
                    pBACe3.6 vector at EcoRI sites"
BASE COUNT      144 a     89 c     89 g    147 t      1 others
ORIGIN


  Query Match             68.3%;  Score 19.8;  DB 230;  Length 470;
  Best Local Similarity   60.9%;  Pred. No. 2e+02;
  Matches   14;  Conservative    7;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 gauucuuuuuguaagccccaagg 26
          | ::|:::::| ||||||||||
Db     76 GTTTCTTTTTGAAAGCCCCAAGG 54


RESULT   9
AQ487601
LOCUS       AQ487601      679 bp    DNA             GSS       24-APR-1999
DEFINITION  RPCI-11-242H19.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-242H19
            , DNA sequence.
ACCESSION   AQ487601
VERSION     AQ487601.1  GI:4673475
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 679)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter
            ,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: RPCI-11-242H19.TV
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genet cs (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: SP6
            Class: BAC ends.
FEATURES            Location/Qualifiers
     source         1. .679
                    /organism="Homo sapiens"
                    /db_xref="GDB:7592730"
                    /db_xref="taxon:9606"
                    /clone="RPCI-11-242H19"
                    /clone_lib="RPCI-11"
                    /sex="Male"
                    /cell_type="Lymphocytes"
                    /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
```

```
                    RPCI11 Human Male BAC Library"
BASE COUNT      200 a    146 c    142 g    191 t
ORIGIN


  Query Match             68.3%;  Score 19.8;  DB 229;  Length 679;
  Best Local Similarity   60.9%;  Pred. No. 2.1e+02;
  Matches   14;  Conservative    7;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 gauucuuuuuguaagccccaagg 26
          | ::|:::::| ||||||||||
Db    501 GTTTCTTTTTGGAAGCCCCAAGG 523


RESULT  10
AQ487546
LOCUS       AQ487546      694 bp    DNA             GSS       24-APR-1999
DEFINITION  RPCI-11-242F17.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-242F17
            , DNA sequence.
ACCESSION   AQ487546
VERSION     AQ487546.1  GI:4673420
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 694)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter
            ,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: RPCI-11-242F17.TV
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genet cs (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .694
                     /organism="Homo sapiens"
                     /db_xref="GDB:7592680"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-242F17"
                     /clone_lib="RPCI-11"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
BASE COUNT      202 a    151 c    143 g    198 t
ORIGIN


  Query Match             68.3%;  Score 19.8;  DB 229;  Length 694;
  Best Local Similarity   60.9%;  Pred. No. 2.1e+02;
  Matches   14;  Conservative    7;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 gauucuuuuuguaagccccaagg 26
          | ::|:::::| ||||||||||
Db    512 GTTTCTTTTTGGAAGCCCCAAGG 534


RESULT  11
AQ487597
LOCUS       AQ487597      723 bp    DNA             GSS       24-APR-1999
DEFINITION  RPCI-11-242H15.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-242H15
            , DNA sequence.
ACCESSION   AQ487597
VERSION     AQ487597.1  GI:4673471
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 723)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter
            ,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: RPCI-11-242H15.TV
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genet cs (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .723
                     /organism="Homo sapiens"
                     /db_xref="GDB:7592726"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-242H15"
                     /clone_lib="RPCI-11"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
BASE COUNT      215 a    150 c    153 g    205 t
ORIGIN


  Query Match             68.3%;  Score 19.8;  DB 229;  Length 723;
  Best Local Similarity   60.9%;  Pred. No. 2.1e+02;
  Matches   14;  Conservative    7;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 gauucuuuuuguaagccccaagg 26
          | ::|:::::| ||||||||||
Db    501 GTTTCTTTTTGGAAGCCCCAAGG 523


RESULT  12
CNS00GMB/c
LOCUS       CNS00GMB     1101 bp    DNA             GSS       03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence T7 end of BAC:
            BACR33B22 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL071964
VERSION     AL071964.1  GI:4952148
KEYWORDS    GSS.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
```

```
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .1101
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone_lib="RPCI-98"
                     /clone="BACR33B22"
                     /note="end : T7"
BASE COUNT      272 a    257 c    187 g    293 t     92 others
ORIGIN


  Query Match             68.3%;  Score 19.8;  DB 219;  Length 1101;
  Best Local Similarity   52.0%;  Pred. No. 2.3e+02;
  Matches   13;  Conservative    9;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 aaagauucuuuuuguaagccccaag 25
          |||||:::::::: :|| ||| |||
Db    616 AAAGATTYTTTTTTTAAACCCAAAG 592


RESULT  13
AW661763
LOCUS       AW661763      332 bp    mRNA            EST       06-APR-2000
DEFINITION  hi80e01.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
            IMAGE:2978616 3', mRNA sequence.
ACCESSION   AW661763
VERSION     AW661763.1  GI:7454389
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 332)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco
            High quality sequence stop: 332.
FEATURES             Location/Qualifiers
     source          1. .332
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2978616"
                     /clone_lib="Soares_NFL_T_GBC_S1"
                     /lab_host="DH10B"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Equal amounts of plasmid DNA from three normalized
                     libraries (fetal lung NbHL19W, testis NHT, and B-cell
                     NCI_CGAP_GCB1) were mixed, and ss circles were made in
                     vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR-amplified cDNAs from pools of 5,000 clones made
                     from the same 3 libraries. The pools consisted of
                     I.M.A.G.E. clones 297480-302087, 682632-687239,
                     726408-728711, and 729096-731399. Subtraction by Bento
                     Soares and M. Fatima Bonaldo.  "
BASE COUNT       99 a     64 c     80 g     89 t
ORIGIN


  Query Match             67.6%;  Score 19.6;  DB 119;  Length 332;
  Best Local Similarity   57.7%;  Pred. No. 2.3e+02;
  Matches   15;  Conservative    7;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 aaagauucuuuuuguaagccccaagg 26
          |||||: |:::::: :| ||| |||||
Db    239 AAAGATCCTTTTTTTATGCCACAAGG 264


RESULT  14
AA761822
LOCUS       AA761822      350 bp    mRNA            EST       07-FEB-1998
DEFINITION  nz29a05.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1289168 3',
            mRNA sequence.
ACCESSION   AA761822
VERSION     AA761822.1  GI:2810752
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 350)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
            Ph.D., Gerald Marti, M.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 749    Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 323.
FEATURES             Location/Qualifiers
     source          1. .350
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1289168"
                     /clone_lib="NCI_CGAP_GCB1"
                     /tissue_type="germinal center B cell"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from human tonsillar cells enriched for
                     germinal center B cells by flow sorting (CD20+, IgD-),
                     provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
                     (NCI) and Dr. Gerald Marti (CBER).  cDNA synthesis was
                     primed with a Not I - oligo(dT) primer
                     [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCCTCATTTTTTTTTTTTTTTTT-3'
                     ]. Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M. Fatima Bonaldo."
```

```
BASE COUNT       93 a     69 c     52 g    136 t
ORIGIN


  Query Match             67.6%;  Score 19.6;  DB 11;  Length 350;
  Best Local Similarity   61.5%;  Pred. No. 2.3e+02;
  Matches   16;  Conservative    6;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 aaagauucuuuuuguaagccccaagg 26
          |||||: |::::| ||| ||||| |
Db    192 AAAGATACTTTTTGGAAGGCCCAATG 217


RESULT  15
AQ561664
LOCUS       AQ561664      428 bp    DNA             GSS       29-MAY-1999
DEFINITION  HS_5206_A1_C10_SP6E RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate=782 Col=19 Row=E, DNA sequence.
ACCESSION   AQ561664
VERSION     AQ561664.1  GI:4921135
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 428)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
            Plate: 782  row: E  column: 19
            Seq primer: SP6
            Class: BAC ends
            High quality sequence stop: 428.
FEATURES             Location/Qualifiers
     source          1. .428
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=782 Col=19 Row=E"
                     /clone_lib="RPCI-11 Human Male BAC Library"
                     /sex="male"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     Male blood DNA was isolated from one randomly chosen donor
                     and partially digested with a combination of EcoRI and
                     EcoRI Methylase. Size selected DNA was cloned into the
                     pBACe3.6 vector at EcoRI sites"
BASE COUNT      140 a     95 c     72 g    116 t      5 others
ORIGIN


  Query Match             67.6%;  Score 19.6;  DB 230;  Length 428;
  Best Local Similarity   53.8%;  Pred. No. 2.4e+02;
  Matches   14;  Conservative    8;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 aaagauucuuuuuguaagccccaagg 26
          |||||::|::::: :   ||||||||
Db     61 AAAGATTCTTTTTTTTCCCCCAAGG 86
```

Search completed: October 2, 2001, 04:38:47
Job time: 4069 sec

GenCore version 4.5


OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 05:00:41 ; Search time 1315.38 Seconds
(without alignments)
341.015 Million cell updates/sec

Title: US-09-310-844B-25
Perfect score: 29
Sequence: 1 aaagauucuuuuuguaagccccaagggcu 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: em_ba1:*
17: em_ba2:*
18: em_fun:*
19: em_htgo_hum:*
20: em_htgo_inv:*
21: em_htgo_rod:*
22: em_htg_hum1:*
23: em_htg_hum2:*
24: em_htg_hum3:*
25: em_htg_hum4:*
26: em_htg_hum5:*
27: em_htg_hum6:*
28: em_htg_hum7:*
29: em_htg_hum8:*
30: em_htg_inv1:*
31: em_htg_inv2:*
32: em_htg_other:*
33: em_htg_rod:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_om:*
43: em_or:*
44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pl:*
48: em_ro:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_vi:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_vi1:*
59: gb_vi2:*
60: gb_htg1:*
61: gb_htg2:*
62: gb_htg3:*
63: gb_htg4:*
64: gb_htg5:*
65: gb_htg6:*
66: gb_htg7:*
67: gb_htg8:*
68: gb_htg9:*
69: gb_htg10:*
70: gb_htg11:*
71: gb_htg12:*
72: gb_htg13:*
73: gb_htg14:*
74: gb_htg15:*
75: gb_htg16:*
76: gb_htg17:*
77: gb_htg18:*
78: gb_htg19:*
79: gb_htg20:*
80: gb_htg21:*
81: gb_htg22:*
82: gb_htg23:*
83: gb_htg24:*
84: gb_htg25:*
85: gb_pr1:*
86: gb_pr2:*
87: gb_pr3:*
88: gb_pr4:*
89: gb_pr5:*
90: gb_pr6:*
91: gb_pr7:*
92: gb_pr8:*
93: gb_pr9:*
94: gb_ro1:*
95: gb_ro2:*
96: gb_in4:*
97: gb_pr10:*
98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 29 | 100.0 | 740 | 95 | RATIL2 | M22899 Rat interle |
| 2 | 29 | 100.0 | 822 | 94 | MMIL03 | X01665 Mouse inter |
| 3 | 29 | 100.0 | 825 | 10 | E00849 | E00849 Nucleic aci |
| 4 | 29 | 100.0 | 825 | 10 | I01928 | I01928 Sequence 1 |
| 5 | 29 | 100.0 | 825 | 94 | MUSIL2T | K02292 Mouse inter |
| 6 | 29 | 100.0 | 868 | 94 | MMINTLK2 | X73040 M.musculus |
| 7 | 29 | 100.0 | 870 | 94 | MUSIL2A03 | M16762 Mouse inter |
| 8 | 29 | 100.0 | 939 | 45 | E10514 | E10514 cDNA encodi |

| | | | | | | |
|---|---|---|---|---|---|---|
| 9 | 29 | 100.0 | 939 | 94 | MMIL04 | X01772 Mouse mRNA |
| 10 | 29 | 100.0 | 24900 | 94 | AF195956 | AF195956 Mus muscu |
| 11 | 24.8 | 85.5 | 764 | 7 | BOVIL2 | M12791 Bovine inte |
| 12 | 24.8 | 85.5 | 773 | 7 | CFU28141 | U28141 Canis famil |
| 13 | 23.2 | 80.0 | 209 | 10 | I82323 | I82323 Sequence 8 |
| 14 | 23.2 | 80.0 | 292 | 54 | G06364 | G06364 human STS W |
| 15 | 23.2 | 80.0 | 689 | 9 | A17717 | A17717 Plasmid pTG |
| 16 | 23.2 | 80.0 | 698 | 9 | A02159 | A02159 Synthetic g |
| 17 | 23.2 | 80.0 | 698 | 9 | A04681 | A04681 Synthetic g |
| 18 | 23.2 | 80.0 | 698 | 9 | A14089 | A14089 Synthetic D |
| 19 | 23.2 | 80.0 | 722 | 10 | I03416 | I03416 Sequence 2 |
| 20 | 23.2 | 80.0 | 730 | 91 | GIBIL2 | K03174 Ape (gibbon |
| 21 | 23.2 | 80.0 | 750 | 10 | I01391 | I01391 Sequence 3 |
| 22 | 23.2 | 80.0 | 769 | 10 | E01460 | E01460 cDNA encodi |
| 23 | 23.2 | 80.0 | 769 | 10 | E02201 | E02201 DNA encodin |
| 24 | 23.2 | 80.0 | 782 | 10 | I01101 | I01101 Sequence 2 |
| 25 | 23.2 | 80.0 | 784 | 9 | A06879 | A06879 Artificial |
| 26 | 23.2 | 80.0 | 784 | 9 | A14095 | A14095 Synthetic D |
| 27 | 23.2 | 80.0 | 784 | 9 | A17732 | A17732 Plasmid pTG |
| 28 | 23.2 | 80.0 | 784 | 10 | E00250 | E00250 DNA coding |
| 29 | 23.2 | 80.0 | 784 | 10 | I08366 | I08366 Sequence 2 |
| 30 | 23.2 | 80.0 | 788 | 10 | I07944 | I07944 Sequence 1 |
| 31 | 23.2 | 80.0 | 790 | 10 | I04498 | I04498 Sequence 1 |
| 32 | 23.2 | 80.0 | 790 | 10 | I08377 | I08377 Sequence 1 |
| 33 | 23.2 | 80.0 | 794 | 10 | E00210 | E00210 cDNA encodi |
| 34 | 23.2 | 80.0 | 794 | 10 | E00211 | E00211 cDNA encodi |
| 35 | 23.2 | 80.0 | 794 | 10 | E02011 | E02011 cDNA encodi |
| 36 | 23.2 | 80.0 | 794 | 10 | E02018 | E02018 cDNA encodi |
| 37 | 23.2 | 80.0 | 795 | 10 | E00214 | E00214 cDNA encodi |
| 38 | 23.2 | 80.0 | 795 | 10 | E00267 | E00267 DNA sequenc |
| 39 | 23.2 | 80.0 | 795 | 10 | E00271 | E00271 DNA sequenc |
| 40 | 23.2 | 80.0 | 801 | 9 | A04961 | A04961 Artificial |
| 41 | 23.2 | 80.0 | 801 | 9 | A06759 | A06759 Artificial |
| 42 | 23.2 | 80.0 | 801 | 10 | E00336 | E00336 Human inter |
| 43 | 23.2 | 80.0 | 801 | 10 | I01197 | I01197 Sequence 1 |
| 44 | 23.2 | 80.0 | 801 | 10 | I04318 | I04318 Sequence 3 |
| 45 | 23.2 | 80.0 | 801 | 93 | HSIL02 | V00564 Human mRNA |

ALIGNMENTS

```
RESULT   1
RATIL2
LOCUS       RATIL2        740 bp    mRNA            ROD       27-APR-1993
DEFINITION  Rat interleukin 2 mRNA, complete cds.
ACCESSION   M22899
VERSION     M22899.1  GI:204909
KEYWORDS    interleukin 2.
SOURCE      Rat activated T-lymphocytes, cDNA to mRNA, clone pRIL-2.8.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 740)
  AUTHORS   McKnight,A.J., Mason,D.W. and Barclay,A.N.
  TITLE     Sequence of rat interleukin 2 and anomalous binding of a mouse
            interleukin 2 cDNA probe to rat MHC class II-associated invariant
            chain mRNA
  JOURNAL   Immunogenetics 30, 145-147 (1989)
  MEDLINE   89339608
COMMENT     Draft entry and computer-readable sequence for [1] kindly submitted
            by A.J.McKnight, 08-MAR-1989.
FEATURES             Location/Qualifiers
     source          1. .740
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
     CDS             15. .482
                     /note="interleukin 2 precursor"
                     /codon_start=1
                     /protein_id="AAA41427.1"
                     /db_xref="GI:204910"
                     /translation="MYSMQLASCVALTLVLLVNSAPTSSPAKETQQHLEQLLLDLQVL
                     LRGIDNYKNLKLPMMLTFKFYLPKQATELKHLQCLENELGALQRVLDLTQSKSFHLED
                     AGNFISNIRVTVVKLKGSENKFECQFDDEPATVVEFLRRWIAICQSIISTMTQ"
     sig_peptide     15. .74
                     /note="interleukin 2 signal peptide"
     mat_peptide     75. .479
                     /note="interleukin 2"
     polyA_signal    713. .718
                     /note="putative"
BASE COUNT      228 a    141 c    143 g    228 t
ORIGIN
```

```
  Query Match             100.0%;  Score 29;  DB 95;  Length 740;
  Best Local Similarity   69.0%;  Pred. No. 0.0039;
  Matches   20;  Conservative    9;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 aaagauucuuuuuguaagccccaagggcu 29
          |||||::|:::::|:||||||||||||||:
Db    612 AAAGATTCTTTTTGTAAGCCCCAAGGGCT 640
```

```
RESULT   2
MMIL03
LOCUS       MMIL03        822 bp    DNA             ROD       07-NOV-1985
DEFINITION  Mouse interleukin-2 gene exon 4 sequence.
ACCESSION   X01665
VERSION     X01665.1  GI:52661
KEYWORDS    interleukin.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 822)
  AUTHORS   Fuse,A., Fujita,T., Yasumitsu,H., Kashima,N., Hasegawa,K. and
            Taniguchi,T.
  TITLE     Organization and structure of the mouse interleukin-2 gene
  JOURNAL   Nucleic Acids Res. 12 (24), 9323-9331 (1984)
  MEDLINE   85087940
FEATURES             Location/Qualifiers
     source          1. .822
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     intron          <1. .246
                     /note="intron III"
     mRNA            247. .741
                     /note="exon 4"
     CDS             247. .357
                     /note="interleukin-2 part 4; Protein sequence is in
                     conflict with the conceptual translation"
                     /codon_start=1
                     /protein_id="CAA25825.1"
                     /db_xref="GI:1334109"
                     /db_xref="MGD:MGI:96548"
                     /db_xref="SWISS-PROT:P04351"
                     /translation="GSDNTFECQFDDESATVVDFLRRWIAFCQSIISTSPQ"
     misc_feature    596. .601
                     /note="polyA signal"
     misc_feature    727. .732
                     /note="polyA signal"
BASE COUNT      273 a    136 c    126 g    287 t
ORIGIN
```

```
  Query Match             100.0%;  Score 29;  DB 94;  Length 822;
  Best Local Similarity   69.0%;  Pred. No. 0.004;
  Matches   20;  Conservative    9;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 aaagauucuuuuuguaagccccaagggcu 29
          |||||::|:::::|:||||||||||||||:
Db    490 AAAGATTCTTTTTGTAAGCCCCAAGGGCT 518
```

```
RESULT   3
```

```
E00849
LOCUS       E00849        825 bp    RNA             PAT       29-SEP-1997
DEFINITION  Nucleic acid sequence of cDNA encoding a polypeptide possessing
            murine IL-2 activity.
ACCESSION   E00849
VERSION     E00849.1  GI:2169110
KEYWORDS    JP 1986119197-A/1.
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 825)
  AUTHORS   Furanku,D.R., Yokota,T. and Arai,K. .
  TITLE     CDNA CLONE FOR CODING POLYPEPTIDE SHOWING RAT INTERLEUKIN-2
            ACTIVITY
  JOURNAL   Patent: JP 1986119197-A 1 06-JUN-1986;
            SHIERINGU BAIOTETSUKU CORP
COMMENT     OS   Human {Homo sapiens}
            PN   JP 1986119197-A/1
            PD   06-JUN-1986
            PF   04-OCT-1985 JP 1985221684
            PR   05-OCT-1984 US 84    658183
            PI   FURANKU DON RII, YOKOTA TAKASHI, ARAI KENICHI PC
            C12P21/02,C07K13/00,C07K15/04,C12N1/00,C12N5/00,C12N15/00, PC
            G01N33/50//
            PC   G01N33/53,G01N33/577,(C12P21/02,C12R1:19),(C12P21/02, PC
            C12R1:865),(C12P21/02,
            PC   C12R1:91),(C12N1/00,C12R1:19),(C12N1/00,C12R1:865),(C12N5/00,
            PC   C12R1:91);
            CC   strandedness: Double;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No;
            CC   *source: tissue_type=Blood;
            CC   *source: cell_type=T-cell;
            CC   *source: cell_line=LB2-1 cell;
            CC   *source: clone=LB2-1 clone;
            FH   Key             Location/Qualifiers
            FH
            FT   CDS             49. .558
            FT                   /product='polypeptide possessing murine IL-2
            FT                   activity'.
FEATURES             Location/Qualifiers
     source          1. .825
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT      252 a    179 c    155 g    239 t
ORIGIN
```

```
  Query Match             100.0%;  Score 29;  DB 10;  Length 825;
  Best Local Similarity   69.0%;  Pred. No. 0.004;
  Matches   20;  Conservative    9;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 aaagauucuuuuuguaagccccaagggcu 29
          |||||::|:::::|:|||||||||||||:
Db    688 AAAGATTCTTTTTGTAAGCCCCAAGGGCT 716
```

```
RESULT   4
I01928
LOCUS       I01928        825 bp ss-DNA             PAT       21-MAY-1993
DEFINITION  Sequence 1 from Patent US 4798789.
ACCESSION   I01928
VERSION     I01928.1  GI:269761
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 825)
  AUTHORS   Lee,F.D., Yokota,T. and Arai,K.
  TITLE     cDNA clones coding for polypeptides exhibiting murine interleukin-2
            activity
  JOURNAL   Patent: US 4798789-A 1 17-JAN-1989;
            DNAX Research Institute of Molecular and Cellular Biology, Inc.;
            Palo Alto, CA
FEATURES             Location/Qualifiers
     source          1. .825
                     /organism="unknown"
BASE COUNT      252 a    179 c    155 g    239 t
ORIGIN
```

```
  Query Match             100.0%;  Score 29;  DB 10;  Length 825;
  Best Local Similarity   69.0%;  Pred. No. 0.004;
  Matches   20;  Conservative    9;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 aaagauucuuuuuguaagccccaagggcu 29
          |||||::|:::::|:|||||||||||||:
Db    688 AAAGATTCTTTTTGTAAGCCCCAAGGGCT 716
```

```
RESULT   5
MUSIL2T
LOCUS       MUSIL2T       825 bp    mRNA            ROD       12-JUN-1993
DEFINITION  Mouse interleukin-2 mRNA, complete cds.
ACCESSION   K02292
VERSION     K02292.1  GI:198330
KEYWORDS    T-cell growth factor; interleukin; interleukin 2.
SOURCE      Mouse (T-cell line LB2-1, from T-cell line C57BL/6), cDNA to mRNA,
            clone MT-1.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 825)
  AUTHORS   Yokota,T., Arai,N., Lee,F., Rennick,D., Mosmann,T. and Arai,K.-I.
  TITLE     Use of a cDNA expression vector for isolation of mouse interleukin
            2 cDNA clones: Expression of T-cell growth-factor activity after
            transfection of monkey cells
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 82, 68-72 (1985)
  MEDLINE   85113172
FEATURES             Location/Qualifiers
     source          1. .825
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     sig_peptide     49. .108
                     /note="interleukin 2 signal peptide"
     CDS             49. .558
                     /note="interleukin 2 prepeptide"
                     /codon_start=1
                     /protein_id="AAA39289.1"
                     /db_xref="GI:309404"
                     /translation="MYSMQLASCVTLTLVLLVNSAPTSSSTSSSTAEAQQQQQQQQQQ
                     QQHLEQLLMDLQELLSRMENYRNLKLPRMLTFKFYLPKQATELKDLQCLEDELGPLRH
                     VLDLTQSKSFQLEDAENFISNIRVTVVKLKGSDNTFECQFDDESATVVDFLRRWIAFC
                     QSIISTSPQ"
     mat_peptide     109. .555
                     /note="interleukin 2"
BASE COUNT      252 a    179 c    155 g    239 t
ORIGIN      45 bp upstream of PstI site.
```

```
  Query Match             100.0%;  Score 29;  DB 94;  Length 825;
  Best Local Similarity   69.0%;  Pred. No. 0.004;
  Matches   20;  Conservative    9;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 aaagauucuuuuuguaagccccaagggcu 29
          |||||::|:::::|:|||||||||||||:
Db    688 AAAGATTCTTTTTGTAAGCCCCAAGGGCT 716
```

```
RESULT   6
MMINTLK2
LOCUS       MMINTLK2      868 bp    mRNA            ROD       17-FEB-1997
DEFINITION  M.musculus mRNA for interleukin-2.
ACCESSION   X73040
VERSION     X73040.1  GI:397468
```

```
KEYWORDS    interleukin 2.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 868)
  AUTHORS   Todd,J.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-APR-1993) J.A. Todd, Nuffield Dept of Surgery,
            University of Oxford, John Radcliffe Hospital, Headington, Oxford,
            OX3 9DU, UK
REFERENCE   2  (bases 1 to 868)
  AUTHORS   Ghosh,S., Palmer,S.M., Rodrigues,N.R., Cordell,H.J., Hearne,C.M.,
            Cornall,R.J., Prins,J.B., McShane,P., Lathrop,G.M., Peterson,L.B.,
            Wicker,L.S. and Todd,J.A.
  TITLE     Polygenic control of autoimmune diabetes in nonobese diabetic mice
  JOURNAL   Nat. Genet. 4 (4), 404-409 (1993)
  MEDLINE   94004970
REFERENCE   3  (bases 1 to 868)
  AUTHORS   Kashima,N., Nishi-Takaoka,C., Fujita,T., Taki,S., Yamada,G.,
            Hamuro,J. and Taniguchi,T.
  TITLE     Unique structure of murine interleukin-2 as deduced from cloned
            cDNAs
  JOURNAL   Nature 313 (6001), 402-404 (1985)
  MEDLINE   85111148
COMMENT     Related sequence: X01772.
FEATURES             Location/Qualifiers
     source          1. .868
                     /organism="Mus musculus"
                     /strain="NOD"
                     /db_xref="taxon:10090"
     mRNA            <1. .>868
                     /product="interleukin-2"
     conflict        81
                     /citation=[3]
                     /replace="t"
     conflict        123
                     /citation=[3]
                     /replace="t"
     conflict        133. .146
                     /note="insertion in NOD"
                     /citation=[3]
                     /replace="gc"
     conflict        161. .162
                     /note="deletion of trinuceotide repeat (cag) in NOD"
                     /citation=[3]
                     /replace="acagcagcagcagc"
BASE COUNT      265 a    190 c    163 g    250 t
ORIGIN
```

```
  Query Match             100.0%;  Score 29;  DB 94;  Length 868;
  Best Local Similarity   69.0%;  Pred. No. 0.004;
  Matches   20;  Conservative    9;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 aaagauucuuuuuguaagccccaagggcu 29
          ||||||::|:::::|:|||||||||||||:
Db    687 AAAGATTCTTTTTGTAAGCCCCAAGGGCT 715
```

```
RESULT   7
MUSIL2A03
LOCUS       MUSIL2A03     870 bp    DNA             ROD       03-JUN-1994
DEFINITION  Mouse interleukin 2 (IL-2) gene, exon 4.
ACCESSION   M16762
VERSION     M16762.1  GI:198306
KEYWORDS    T-cell growth factor; interleukin 2.
SEGMENT     3 of 3
SOURCE      Mouse DNA, clones pATgMIL2-A and pBRgMIL2-3.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 870)
  AUTHORS   Degrave,W.M., Simons,G., Devos,R., Plaetinck,G., Remaut,E.,
            Tavernier,J. and Fiers,W.
  TITLE     Cloning and structure of a mouse interleukin-2 chromosomal gene
  JOURNAL   Mol. Biol. Rep. 11, 57-61 (1986)
  MEDLINE   86118396
FEATURES             Location/Qualifiers
     source          1. .870
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /map="3"
     gene            join(M16760.1:213. .648,M16761.1:1. .559,1. .870)
                     /gene="Il2"
     CDS             join(M16760.1:260. .448,M16760.1:548. .607,
                     M16761.1:148. .294,70. .183)
                     /gene="Il2"
                     /note="precursor"
                     /codon_start=1
                     /product="interleukin 2"
                     /protein_id="AAA39281.1"
                     /db_xref="GI:387384"
                     /translation="MYSMQLASCVTLTLVLLVNSAPTSSSTSSSTAEAQQQQQQQQQQ
                     QQHLEQLLMDLQELLSRMENYRNLKLPRMLTFKFYLPKQATELKDLQCLEDELGPLRH
                     VLDLTESKSFQLEDAENFISNIRVTVVKLKGSDNTFECQFDDESATVVDFLRRWIAFC
                     QSIISTSPQ"
     mat_peptide     join(M16760.1:320. .448,M16760.1:548. .607,
                     M16761.1:148. .294,70. .180)
                     /gene="Il2"
                     /note="precursor"
                     /product="interleukin 2"
     intron          <1. .69
                     /gene="Il2"
                     /number=3
     exon            70. .>870
                     /gene="Il2"
                     /number=4
BASE COUNT      284 a    150 c    135 g    301 t
ORIGIN      1 bp upstream of HindIII site; about 1900 bp after segment 1.
```

```
  Query Match             100.0%;  Score 29;  DB 94;  Length 870;
  Best Local Similarity   69.0%;  Pred. No. 0.004;
  Matches   20;  Conservative    9;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 aaagauucuuuuuguaagccccaagggcu 29
          ||||||::|:::::|:|||||||||||||:
Db    313 AAAGATTCTTTTTGTAAGCCCCAAGGGCT 341
```

```
RESULT   8
E10514
ID   E10514     standard; RNA; ROD; 939 BP.
XX
AC   E10514;
XX
SV   E10514.1
XX
DT   08-OCT-1997 (Rel. 52, Created)
DT   02-SEP-2000 (Rel. 65, Last updated, Version 2)
XX
DE   cDNA encoding mouse interleukin-2.
XX
KW   JP 1996000272-A/1.
XX
OS   Mus sp.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
XX
RN   [1]
RP   1-939
RA   Kashima S., Hamuro J., Taniguchi K., Sugano H.;
RT   "DNA FRAGMENT CONTAINING GENE CODING MOUSE INTERLEUIKIN-2, RECOMBINANT
RT   VECTOR HAVING THE SAME, AND TRANSFORMANT HOLDING THE";
RL   Patent number JP1996000272-A/1, 09-JAN-1996.
```

```
RL   AJINOMOTO CO INC, JAPAN FOUND CANCER RES.
XX
CC   OS   Mus sp. (mouse)
CC   PN   JP 1996000272-A/1
CC   PD   09-JAN-1996
CC   PF   26-JUL-1988 JP 1995180603
CC   PI   KASHIMA SHINICHI, HAMURO JUNJI, TANIGUCHI KOREAKI,
CC   PI   SUGANO HARUO
CC   PC   C12N15/09,C07H21/04,C12N1/21,C12N5/10,C12P21/02//A61K35/12,
CC   PC   A61K35/74,
CC   PC   A61K38/00,C07K14/55,(C12N1/21,C12R1:19),(C12P21/02,C12R1:19),
CC   PC   (C12P21/02,
CC   PC   C12R1:91);
CC   CC   strandedness: Double;
CC   CC   topology: Linear;
CC   CC   hypothetical: No;
CC   CC   anti-sense: No;
CC   FH   Key             Location/Qualifiers
CC   FH
CC   FT   source          1. .939
CC   FT                   /organism="Mus sp."
CC   FT                   /cell_line="LBRM-33"
CC   FT                   /tissue_type="lymphoma"
CC   FT   5'UTR           1. .47
CC   FT   sig_peptide     48. .110
CC   FT                   /product="signal peptide of mouse
CC   FT                   interleukin-2"
CC   FT   mat_peptide     111. .554
CC   FT                   /product="mouse interleukin-2"
CC   FT   CDS             48. .557
CC   FT                   /product="mouse interleukin-2"
CC   FT   3'UTR           558. .939
XX
FH   Key             Location/Qualifiers
FH
FT   source          1. .939
FT                   /db_xref="taxon:10095"
FT                   /organism="Mus sp."
XX
SQ   Sequence 939 BP; 287 A; 202 C; 173 G; 277 T; 0 other;


  Query Match             100.0%;  Score 29;  DB 45;  Length 939;
  Best Local Similarity   69.0%;  Pred. No. 0.004;
  Matches   20;  Conservative    9;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 aaagauucuuuuuguaagccccaagggcu 29
          |||||::|:::::|:|||||||||||||:
Db    687 AAAGATTCTTTTTGTAAGCCCCAAGGGCT 715


RESULT   9
MMIL04
LOCUS       MMIL04        939 bp    mRNA            ROD       30-MAR-1995
DEFINITION  Mouse mRNA for interleukin-2 (IL-2).
ACCESSION   X01772 K02797
VERSION     X01772.1  GI:52663
KEYWORDS    interleukin; signal peptide; tandem repeat.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 939)
  AUTHORS   Kashima,N., Nishi-Takaoka,C., Fujita,T., Taki,S., Yamada,G.,
            Hamuro,J. and Taniguchi,T.
  TITLE     Unique structure of murine interleukin-2 as deduced from cloned
            cDNAs
  JOURNAL   Nature 313 (6001), 402-404 (1985)
  MEDLINE   85111148
COMMENT     Data kindly reviewed (19-FEB-1986) by G. Yamada.
FEATURES             Location/Qualifiers
     source          1. .939
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     CDS             48. .557
                     /codon_start=1
                     /product="interleukin-2"
                     /protein_id="CAA25909.1"
                     /db_xref="GI:758159"
                     /db_xref="MGD:MGI:96548"
                     /db_xref="SWISS-PROT:P04351"
                     /translation="MYSMQLASCVTLTLVLLVNSAPTSSSTSSSTAEAQQQQQQQQQQ
                     QQHLEQLLMDLQELLSRMENYRNLKLPRMLTFKFYLPKQATELKDLQCLEDELGPLRH
                     VLDLTQSKSFQLEDAENFISNIRVTVVKLKGSDNTFECQFDDESATVVDFLRRWIAFC
                     QSIISTSPQ"
     sig_peptide     48. .107
                     /product="signal peptide (aa -21 to -1)"
     mat_peptide     108. .554
                     /product="interleukin-2 (aa 1-133)"
     repeat_region   150. .185
                     /note="12 tandem repeats of CAG"
BASE COUNT      287 a    202 c    173 g    277 t
ORIGIN


  Query Match             100.0%;  Score 29;  DB 94;  Length 939;
  Best Local Similarity   69.0%;  Pred. No. 0.004;
  Matches   20;  Conservative    9;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 aaagauucuuuuuguaagccccaagggcu 29
          |||||::|:::::|:|||||||||||||:
Db    687 AAAGATTCTTTTTGTAAGCCCCAAGGGCT 715


RESULT  10
AF195956
LOCUS       AF195956    24900 bp    DNA             ROD       04-MAY-2000
DEFINITION  Mus musculus interleukin-2 (Il2) gene, complete cds; and testis
            nuclear RNA-binding protein (Tenr) gene, partial cds.
ACCESSION   AF195956
VERSION     AF195956.1  GI:6941952
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 24900)
  AUTHORS   Lyons,P.A., Armitage,N., Argentina,F., Denny,P., Hill,N.J.,
            Lord,C.J., Wilusz,M.B., Peterson,L.B., Wicker,L.S. and Todd,J.A.
  TITLE     Congenic mapping of the type 1 diabetes locus, idd3, to a 780-kb
            region of mouse chromosome 3: identification of a candidate segment
            of ancestral DNA by haplotype mapping
  JOURNAL   Genome Res. 10 (4), 446-453 (2000)
  MEDLINE   20243845
   PUBMED   10779485
REFERENCE   2  (bases 1 to 24900)
  AUTHORS   Lyons,P.A., Armitage,N., Lord,C.J., Denny,P., Hill,N.J.,
            Podolin,P.L., Peterson,L.B., Wicker,L.S. and Todd,J.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-OCT-1999) Medical Genetics, University of Cambridge,
            Cambridge Institute for Medical Research, Addenbrooke's Hospital,
            Cambridge CB2 2XY, UK
FEATURES             Location/Qualifiers
     source          1. .24900
                     /organism="Mus musculus"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /chromosome="3"
                     /clone="mC1h5"
     repeat_region   complement(385. .509)
                     /note="SINE"
                     /rpt_family="B1"
                     /rpt_type=dispersed
     repeat_region   2786. .2853
                     /note="SINE"
                     /rpt_family="MIR"
```

```
                     /rpt_type=dispersed
     repeat_region   3411. .3595
                     /note="SINE"
                     /rpt_family="B4"
                     /rpt_type=dispersed
     mRNA            join(<4562. .4750,4850. .4909,7333. .7479,9254. .9749)
                     /gene="Il2"
                     /product="Il2"
                     /evidence=not_experimental
     gene            <4562. .9749
                     /gene="Il2"
                     /evidence=not_experimental
     CDS             join(4562. .4750,4850. .4909,7333. .7479,9254. .9367)
                     /gene="Il2"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="interleukin-2"
                     /protein_id="AAF32272.1"
                     /db_xref="GI:6941953"
                     /translation="MYSMQLASCVTLTLVLLVNSAPTSSSTSSSTAEAQQQQQQQQQQ
                     QQHLEQLLMDLQELLSRMENYRNLKLPRMLTFKFYLPKQATELKDLQCLEDELGPLRH
                     VLDLTQSKSFQLEDAENFISNIRVTVVKLKGSDNTFECQFDDESATVVDFLRRWIAFC
                     QSIISTSPQ"
     repeat_region   4664. .4699
                     /note="microsatellite D3Mit21"
                     /rpt_type=tandem
                     /rpt_unit=cag
     repeat_region   complement(6642. .6794)
                     /note="SINE"
                     /rpt_family="B4"
                     /rpt_type=dispersed
     repeat_region   complement(8109. .8242)
                     /note="SINE"
                     /rpt_family="B1"
                     /rpt_type=dispersed
     repeat_region   10250. .10391
                     /note="microsatellite D3Nds6; sequence deposited at
                     GenBank Accession Number X58264"
                     /rpt_type=tandem
     repeat_region   11958. .12060
                     /note="SINE"
                     /rpt_family="B4"
                     /rpt_type=dispersed
     repeat_region   12305. .12639
                     /note="LINE"
                     /rpt_family="L1"
                     /rpt_type=dispersed
     repeat_region   13781. .13916
                     /note="SINE"
                     /rpt_family="B1"
                     /rpt_type=dispersed
     repeat_region   14163. .17181
                     /note="LINE"
                     /rpt_family="L1"
                     /rpt_type=dispersed
     repeat_region   17421. .17595
                     /note="SINE"
                     /rpt_family="B3"
                     /rpt_type=dispersed
     mRNA            complement(join(18965. .19123,24242. .>24374))
                     /gene="Tenr"
                     /product="testis nuclear RNA-binding protein"
                     /evidence=not_experimental
     gene            complement(18965. .>24900)
                     /gene="Tenr"
                     /evidence=not_experimental
     CDS             complement(join(19010. .19123,24242. .>24374))
                     /gene="Tenr"
                     /note="TENR"
                     /codon_start=2
                     /evidence=not_experimental
                     /product="testis nuclear RNA-binding protein"
                     /protein_id="AAF32273.1"
                     /db_xref="GI:6941954"
                     /translation="SPFKSGLSMASRLCKAAMLSRFNLLAKEAKTDDLLEARTYHAAK
                     CMSGPYQEAKALLKAYLQQHGYGSWIVKSPCIEQFSM"
     repeat_region   complement(19651. .20127)
                     /note="LINE"
                     /rpt_family="L1"
                     /rpt_type=dispersed
     repeat_region   complement(20958. .21142)
                     /note="LINE"
                     /rpt_family="L1"
                     /rpt_type=dispersed
     repeat_region   complement(22450. .22619)
                     /note="LINE"
                     /rpt_family="L1"
                     /rpt_type=dispersed
     repeat_region   complement(23262. .23446)
                     /note="LINE"
                     /rpt_family="L1"
                     /rpt_type=dispersed
     repeat_region   complement(24776. .24898)
                     /note="SINE"
                     /rpt_family="B2"
                     /rpt_type=dispersed
BASE COUNT     7948 a    4720 c    4896 g    7278 t      58 others
ORIGIN


  Query Match             100.0%;  Score 29;  DB 94;  Length 24900;
  Best Local Similarity    69.0%;  Pred. No. 0.0068;
  Matches   20;  Conservative    9;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 aaagauucuuuuuguaagccccaagggcu 29
           |||||::|:::::|:||||||||||||||:
Db    9497 AAAGATTCTTTTTGTAAGCCCCAAGGGCT 9525


RESULT  11
BOVIL2
LOCUS       BOVIL2        764 bp    mRNA            MAM       27-APR-1993
DEFINITION  Bovine interleukin 2 (IL-2) mRNA, complete cds.
ACCESSION   M12791
VERSION     M12791.1  GI:163204
KEYWORDS    interleukin 2.
SOURCE      Bovine lymph node, cDNA to mRNA, clone pBIL2-4.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 764)
  AUTHORS   Cerretti,D.P., McKereghan,K., Larsen,A., Cantrell,M.A.,
            Anderson,D., Gillis,S., Cosman,D. and Baker,P.E.
  TITLE     Cloning, sequence, and expression of bovine interleukin 2
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 83, 3223-3227 (1986)
  MEDLINE   86205869
COMMENT     Draft entry and clean copy sequence for [1] kindly provided by
            D.Cerretti, 12-AUG-1986.
            There is probably only one copy of the interleukin 2 gene in the
            bovine genome.
FEATURES             Location/Qualifiers
     source          1. .764
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /tissue_type="lymph"
     mRNA            <1. .764
                     /gene="IL-2"
     gene            1. .764
                     /gene="IL-2"
     sig_peptide     18. .77
                     /gene="IL-2"
     CDS             18. .485
                     /gene="IL-2"
                     /note="prepeptide"
                     /codon_start=1
```

```
                     /product="interleukin 2"
                     /protein_id="AAA30586.1"
                     /db_xref="GI:163205"
                     /translation="MYKIQLLSCIALTLALVANGAPTSSSTGNTMKEVKSLLLDLQLL
                     LEKVKNPENLKLSRMHTFDFYVPKVNATELKHLKCLLEELKLLEEVLNLAPSKNLNPR
                     EIKDSMDNIKRIVLELQGSETRFTCEYDDATVNAVEFLNKWITFCQSIYSTMT"
     mat_peptide     78. .482
                     /gene="IL-2"
                     /product="interleukin 2"
BASE COUNT      257 a    133 c    123 g    251 t
ORIGIN      80 bp upstream of HgiAI site.


  Query Match             85.5%;  Score 24.8;  DB 7;  Length 764;
  Best Local Similarity   60.7%;  Pred. No. 0.38;
  Matches   17;  Conservative    9;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 aagauucuuuuuguaagccccaagggcu 29
          ||||::|:::::|:|||||| | ||||:
Db    620 AAGATTCTTTTTGTAAGCCCTACGGGCT 647


RESULT  12
CFU28141
LOCUS       CFU28141      773 bp    mRNA            MAM       29-JUN-1995
DEFINITION  Canis familiaris interleukin-2 mRNA, complete cds.
ACCESSION   U28141
VERSION     U28141.1  GI:881935
KEYWORDS    .
SOURCE      dog.
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1  (bases 1 to 773)
  AUTHORS   Somberg,R.L., Tipold,A., Henthorn,P.S. and Felsburg,P.J.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 773)
  AUTHORS   Somberg,R.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-MAY-1995) Richard L. Somberg, Clinical Studies,
            University of Pennsylvania, 3850 Spruce Street, Philadelphia, PA
            19104, USA
FEATURES             Location/Qualifiers
     source          1. .773
                     /organism="Canis familiaris"
                     /db_xref="taxon:9615"
     CDS             39. .506
                     /codon_start=1
                     /product="interleukin-2"
                     /protein_id="AAA68969.1"
                     /db_xref="GI:881936"
                     /translation="MYKMQLLSCIALTLVLVANSAPITSSSTKETEQQMEQLLLDLQL
                     LLNGVNNYENPQLSRMLTFKFYTPKKATEFTHLQCLAEELKNLEEVLGLPQSKNVHLT
                     DTKELISNMNVTLLKLKGSETSYNCEYDDETATITEFLNKWITFCQSIFSTLT"
     polyA_site      773
                     /note="38 A nucleotides"
BASE COUNT      265 a    145 c    118 g    245 t
ORIGIN


  Query Match             85.5%;  Score 24.8;  DB 7;  Length 773;
  Best Local Similarity   60.7%;  Pred. No. 0.38;
  Matches   17;  Conservative    9;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 aagauucuuuuuguaagccccaagggcu 29
          ||||::|:::::|:|||||| | ||||:
Db    641 AAGATTCTTTTTGTAAGCCCTAGGGGCT 668


RESULT  13
I82323
LOCUS       I82323        209 bp    DNA             PAT       10-JUN-1998
DEFINITION  Sequence 8 from patent US 5712126.
ACCESSION   I82323
VERSION     I82323.1  GI:3210620
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 209)
  AUTHORS   Weissman,S.M. and Prashar,Y.
  TITLE     Analysis of gene expression by display of 3-end restriction
            fragments of CDNA
  JOURNAL   Patent: US 5712126-A 8 27-JAN-1998;
FEATURES             Location/Qualifiers
     source          1. .209
                     /organism="unknown"
BASE COUNT       56 a     30 c     33 g     90 t
ORIGIN


  Query Match             80.0%;  Score 23.2;  DB 10;  Length 209;
  Best Local Similarity   57.1%;  Pred. No. 1.7;
  Matches   16;  Conservative    9;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 aagauucuuuuuguaagccccaagggcu 29
          | ||::|:::::|:|||||| | ||||:
Db     34 ATGATTCTTTTTGTAAGCCCTAGGGGCT 61


RESULT  14
G06364
LOCUS       G06364        292 bp    DNA             STS       19-OCT-1995
DEFINITION  human STS WI-7035.
ACCESSION   G06364
VERSION     G06364.1  GI:859609
KEYWORDS    STS sequence; primer; sequence tagged site.
SOURCE      human STSs derived from sequences in dbEST and the Unigene
            collection.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata;
            Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates;
            Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 292)
  AUTHORS   Hudson,T.
  TITLE     Whitehead Institute/MIT Center for Genome Research; Physically
            Mapped ESTs
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Thomas Hudson
            Whitehead Institute/MIT Center for Genome Research
            Whitehead Institute for Biomedical Research
            9 Cambridge Center, Cambridge MA 02142 USA
            Tel: 617 252 1900
            Fax: 617 252 1902
            Email: thudson@genome.wi.mit.edu

            Primer A: TAATTAAGTGCTTCCCACTTAAAAC
            Primer B: ATTTTGGGATAAATAAGTGAAACCA
            STS size: 200
            PCR Profile:
                  Presoak:
                  Denaturation:
                  Annealing: 56 degrees C
                  Polymerization:
                  PCR Cycles: 35
                  Thermal Cycler:
            Protocol:
                Template: 10 ng
                Primer: each 5 pM
                dNTPs: each 4 nM
                Taq Polymerase: 0.025 units/ul
                Total Vol: 20 ul

            Buffer:
```

```
              MgCl2: 1.5 mM
              KCl: 50 mM
              Tris-HCL: 10 mM
              pH: 9.3

          Prepared with primer pairs derived from V00564 -- Unigene.
FEATURES             Location/Qualifiers
     source          1. .292
                     /organism="Homo sapiens"
     STS             1. .200
     primer_bind     1. .25
     primer_bind     complement(176. .200)
BASE COUNT       92 a     32 c     32 g    124 t     12 others
ORIGIN


  Query Match             80.0%;  Score 23.2;  DB 54;  Length 292;
  Best Local Similarity   57.1%;  Pred. No. 1.8;
  Matches   16;  Conservative    9;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 aagauucuuuuuguaagccccaagggcu 29
           | ||::|:::::|:|||||| | ||||:
Db     142 ATGATTCTTTTTGTAAGCCCTAGGGGCT 169


RESULT  15
A17717
LOCUS       A17717        689 bp    mRNA            PAT       17-MAY-1994
DEFINITION  Plasmid pTG 26 mRNA for human IL-2.
ACCESSION   A17717
VERSION     A17717.1  GI:513949
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  Plasmid unidentified
            unclassified.
REFERENCE   1  (bases 1 to 689)
  AUTHORS   Kieny,M.P., Sondermeyer,P. and Lecocq,J.P.
  TITLE     Expression of human IL-2 in mammalian cells by a recombined pox
            virus
  JOURNAL   Patent: EP 0206939-A 1 30-DEC-1986;
            TRANSGENE S.A
FEATURES             Location/Qualifiers
     source          1. .689
                     /organism="unidentified"
                     /plasmid="pTG26"
                     /db_xref="taxon:32644"
     CDS             1. .384
                     /partial
                     /codon_start=1
                     /product="human interleukin 2"
                     /protein_id="CAA01347.1"
                     /db_xref="GI:513950"
                     /translation="TKKTQLQLEHLLLDLQMILNGINNYKNPKLTRMLTFKFYMPKKA
                     TELKHLQCLEEELKPLEEVLNLAQSKNFHLRPRDLISNINVIVLELKGSETTFMCEYA
                     DETATIVEFLNRWITFCQSIISTLT"
BASE COUNT      251 a    117 c     99 g    222 t
ORIGIN


  Query Match             80.0%;  Score 23.2;  DB 9;  Length 689;
  Best Local Similarity   57.1%;  Pred. No. 2.1;
  Matches   16;  Conservative    9;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 aagauucuuuuuguaagccccaagggcu 29
           | ||::|:::::|:|||||| | ||||:
Db     526 ATGATTCTTTTTGTAAGCCCTAGGGGCT 553


Search completed: October  2, 2001, 05:00:42
Job time: 5329 sec
```

GenCore version 4.5


OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 05:03:57 ; Search time 122.48 Seconds
(without alignments)
148.670 Million cell updates/sec

Title: US-09-310-844B-25
Perfect score: 29
Sequence: 1 aaagauucuuuuuguaagccccaagggcu 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

tal number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_0601:*
1: /SIDS1/gcgdata/geneseq/geneseqn/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 29 | 100.0 | 29 | 21 | AAA70829 | Molecular interact |
| 2 | 29 | 100.0 | 29 | 21 | AAA70830 | Molecular interact |
| 3 | 29 | 100.0 | 42 | 21 | AAA71115 | Molecular interact |
| 4 | 29 | 100.0 | 42 | 21 | AAA71116 | Molecular interact |
| 5 | 29 | 100.0 | 42 | 21 | AAA71120 | Molecular interact |
| 6 | 29 | 100.0 | 42 | 21 | AAA71121 | Molecular interact |
| 7 | 29 | 100.0 | 42 | 21 | AAA71128 | Molecular interact |
| 8 | 29 | 100.0 | 42 | 21 | AAA71129 | Molecular interact |
| 9 | 29 | 100.0 | 825 | 7 | AAN60149 | Sequence encoding |
| 10 | 29 | 100.0 | 940 | 7 | AAN60906 | Sequence encodes m |
| 11 | 28 | 96.6 | 45 | 21 | AAA70825 | Molecular interact |
| 12 | 28 | 96.6 | 45 | 21 | AAA70826 | Molecular interact |
| 13 | 28 | 96.6 | 46 | 21 | AAA71088 | Molecular interact |
| 14 | 28 | 96.6 | 46 | 21 | AAA71089 | Molecular interact |
| 15 | 28 | 96.6 | 46 | 21 | AAA71090 | Molecular interact |
| 16 | 28 | 96.6 | 46 | 21 | AAA71105 | Molecular interact |
| 17 | 28 | 96.6 | 46 | 21 | AAA71106 | Molecular interact |
| 18 | 28 | 96.6 | 46 | 21 | AAA71107 | Molecular interact |
| 19 | 24.8 | 85.5 | 42 | 21 | AAA71113 | Molecular interact |
| 20 | 24.8 | 85.5 | 42 | 21 | AAA71118 | Molecular interact |
| 21 | 24.8 | 85.5 | 42 | 21 | AAA71126 | Molecular interact |
| 22 | 23.8 | 82.1 | 46 | 21 | AAA71085 | Molecular interact |
| 23 | 23.8 | 82.1 | 46 | 21 | AAA71103 | Molecular interact |
| 24 | 23.2 | 80.0 | 29 | 21 | AAA70828 | Molecular interact |
| 25 | 23.2 | 80.0 | 42 | 21 | AAA71123 | Molecular interact |
| 26 | 23.2 | 80.0 | 42 | 21 | AAA71131 | Molecular interact |
| 27 | 23.2 | 80.0 | 209 | 18 | AAT61453 | IL-2 cDNA 3'end. |
| 28 | 23.2 | 80.0 | 209 | 21 | AAZ59456 | Interleukin-2 frag |
| 29 | 23.2 | 80.0 | 698 | 6 | AAN50132 | DNA sequence encod |
| 30 | 23.2 | 80.0 | 698 | 7 | AAN60101 | Sequence of human |
| 31 | 23.2 | 80.0 | 722 | 8 | AAN70144 | Sequence encoding |
| 32 | 23.2 | 80.0 | 769 | 11 | AAQ01764 | Human interleukin- |
| 33 | 23.2 | 80.0 | 784 | 5 | AAN40046 | Sequence encoding |
| 34 | 23.2 | 80.0 | 784 | 7 | AAN60787 | Cloned sequence en |
| 35 | 23.2 | 80.0 | 784 | 7 | AAN60102 | Sequence of human |
| 36 | 23.2 | 80.0 | 788 | 10 | AAN90253 | Interleukin-2. Ho |
| 37 | 23.2 | 80.0 | 794 | 5 | AAN40441 | Gene encoding poly |
| 38 | 23.2 | 80.0 | 794 | 10 | AAN90381 | Recombinant human |
| 39 | 23.2 | 80.0 | 800 | 6 | AAN50219 | Sequence encoding |
| 40 | 23.2 | 80.0 | 801 | 4 | AAN30031 | Sequence of interl |
| 41 | 23.2 | 80.0 | 801 | 5 | AAN40057 | Cloned human inter |
| 42 | 23.2 | 80.0 | 801 | 5 | AAN40254 | Sequence encoding |
| 43 | 23.2 | 80.0 | 802 | 6 | AAN50279 | DNA sequence conta |
| 44 | 23.2 | 80.0 | 810 | 7 | AAN60840 | Sequence encoding |
| 45 | 23.2 | 80.0 | 844 | 21 | AAF20965 | Human low adenosin |

ALIGNMENTS

RESULT 1
AAA70829
ID AAA70829 standard; RNA; 29 BP.
XX
AC AAA70829;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #29.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Mus sp.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -
XX

PS Claim 235; Page 235; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 29 BP; 8 A; 6 C; 6 G; 9 U; 0 other;

Query Match 100.0%; Score 29; DB 21; Length 29;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 aaagauucuuuuuguaagccccaagggcu 29
||||||||||||||||||||||||||||||
Db 1 aaagauucuuuuuguaagccccaagggcu 29

RESULT 2
AAA70830
ID AAA70830 standard; RNA; 29 BP.
XX
AC AAA70830;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #30.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Rattus sp.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -
XX

PS Claim 235; Page 235; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 29 BP; 8 A; 6 C; 6 G; 9 U; 0 other;

Query Match 100.0%; Score 29; DB 21; Length 29;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 aaagauucuuuuuguaagccccaagggcu 29
||||||||||||||||||||||||||||||
Db 1 aaagauucuuuuuguaagccccaagggcu 29

RESULT 3
AAA71115
ID AAA71115 standard; RNA; 42 BP.
XX
AC AAA71115;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #191.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -
XX

PS Example 7; Figure 122; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 42 BP; 13 A; 7 C; 7 G; 15 U; 0 other;

Query Match 100.0%; Score 29; DB 21; Length 42;
Best Local Similarity 100.0%; Pred. No. 0.00038;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 aaagauucuuuuuguaagccccaagggcu 29
||||||||||||||||||||||||||||||
Db 4 aaagauucuuuuuguaagccccaagggcu 32

RESULT 4
AAA71116
ID AAA71116 standard; RNA; 42 BP.
XX
AC AAA71116;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #192.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -
XX

PS Example 7; Figure 122; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 42 BP; 13 A; 7 C; 7 G; 15 U; 0 other;

Query Match 100.0%; Score 29; DB 21; Length 42;
Best Local Similarity 100.0%; Pred. No. 0.00038;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 aaagauucuuuuuguaagccccaagggcu 29
||||||||||||||||||||||||||||||
Db 4 aaagauucuuuuuguaagccccaagggcu 32

RESULT 5
AAA71120
ID AAA71120 standard; DNA; 42 BP.
XX
AC AAA71120;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #126.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -
XX

```
PS   Example 7; Figure 125; 405pp; English.
XX
CC   This invention describes a novel method for identifying compounds which
CC   modulate the activity of a target biomolecule. The method uses
CC   3-dimensional representations of the biomolecule and a library of
CC   compounds and comprises (a) identifying at least one molecular
CC   interaction site of the target RNA; (b) generating in silico a virtual
CC   library of compounds predicted or calculated to interact with the
CC   molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC   representations of the target RNA with members of the virtual library of
CC   compounds to generate a hierarchy of the compounds ranked in accordance
CC   with their respective ability to form physical interactions with the
CC   molecular interaction site. The method also describes (1) RNA comprising
CC   a joined sequence of at least 24 nucleotides but not more than 70
CC   nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC   forming a first side of a first double stranded (ds) region; (b) 2
CC   nucleotides forming a first side of an internal loop region; (c) 4
CC   nucleotides forming a first side of a second ds region; (d) 4 or 5
CC   nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC   second side of the second ds region; (f) 4 nucleotides forming a second
CC   side of the internal loop region; and (g) 3 nucleotides forming a second
CC   side of the first ds region; (2) a purified and isolated RNA fragment
CC   comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC   methods and products can be used for identifying agents which modulate
CC   the activity of biomolecules, particularly RNA. Such agents can be used
CC   as pharmaceutical, agricultural or industrial compounds.
XX
SQ   Sequence 42 BP; 13 A; 7 C; 7 G; 15 T; 0 other;


  Query Match             100.0%;  Score 29;  DB 21;  Length 42;
  Best Local Similarity   69.0%;  Pred. No. 0.00038;
  Matches   20;  Conservative    9;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 aaagauucuuuuuguaagccccaagggcu 29
           |||||::|:::::|:||||||||||||:
Db       4 aaagattcttttgtaagccccaagggct 32


RESULT   6
AAA71121
ID   AAA71121 standard; DNA; 42 BP.
XX
AC   AAA71121;
XX
DT   27-APR-2001  (first entry)
XX
DE   Molecular interaction site DNA #127.
XX
KW   Modulator; identification; molecular interaction; virtual library; ss.
XX
OS   Unidentified.
XX
PN   WO9958947-A2.
XX
PD   18-NOV-1999.
XX
PF   12-MAY-1999;   99WO-US10361.
XX
PR   12-MAY-1998;   98US-0076404.
PR   12-MAY-1998;   98US-0085092.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Ecker DJ,  Griffey R,  Crooke ST,  Sampath R,  Swayze E,  Mohan V;
PI   Hofstadler S,  McNeil J;
XX
DR   WPI; 2000-086439/07.
XX
PT   Identifying compounds which modulate activity of target biomolecules,
PT   used to provide compounds which can be used as pharmacological,
PT   agricultural and industrial compounds -
XX
```

```
PS   Example 7; Figure 125; 405pp; English.
XX
CC   This invention describes a novel method for identifying compounds which
CC   modulate the activity of a target biomolecule. The method uses
CC   3-dimensional representations of the biomolecule and a library of
CC   compounds and comprises (a) identifying at least one molecular
CC   interaction site of the target RNA; (b) generating in silico a virtual
CC   library of compounds predicted or calculated to interact with the
CC   molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC   representations of the target RNA with members of the virtual library of
CC   compounds to generate a hierarchy of the compounds ranked in accordance
CC   with their respective ability to form physical interactions with the
CC   molecular interaction site. The method also describes (1) RNA comprising
CC   a joined sequence of at least 24 nucleotides but not more than 70
CC   nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC   forming a first side of a first double stranded (ds) region; (b) 2
CC   nucleotides forming a first side of an internal loop region; (c) 4
CC   nucleotides forming a first side of a second ds region; (d) 4 or 5
CC   nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC   second side of the second ds region; (f) 4 nucleotides forming a second
CC   side of the internal loop region; and (g) 3 nucleotides forming a second
CC   side of the first ds region; (2) a purified and isolated RNA fragment
CC   comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC   methods and products can be used for identifying agents which modulate
CC   the activity of biomolecules, particularly RNA. Such agents can be used
CC   as pharmaceutical, agricultural or industrial compounds.
XX
SQ   Sequence 42 BP; 13 A; 7 C; 7 G; 15 T; 0 other;


  Query Match             100.0%;  Score 29;  DB 21;  Length 42;
  Best Local Similarity   69.0%;  Pred. No. 0.00038;
  Matches   20;  Conservative    9;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 aaagauucuuuuuguaagccccaagggcu 29
           |||||::|:::::|:||||||||||||:
Db       4 aaagattcttttgtaagccccaagggct 32


RESULT   7
AAA71128
ID   AAA71128 standard; RNA; 42 BP.
XX
AC   AAA71128;
XX
DT   27-APR-2001  (first entry)
XX
DE   Molecular interaction site RNA #197.
XX
KW   Modulator; identification; molecular interaction; virtual library; ss.
XX
OS   Unidentified.
XX
PN   WO9958947-A2.
XX
PD   18-NOV-1999.
XX
PF   12-MAY-1999;   99WO-US10361.
XX
PR   12-MAY-1998;   98US-0076404.
PR   12-MAY-1998;   98US-0085092.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Ecker DJ,  Griffey R,  Crooke ST,  Sampath R,  Swayze E,  Mohan V;
PI   Hofstadler S,  McNeil J;
XX
DR   WPI; 2000-086439/07.
XX
PT   Identifying compounds which modulate activity of target biomolecules,
PT   used to provide compounds which can be used as pharmacological,
PT   agricultural and industrial compounds -
XX
```

PS Example 7; Figure 126; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 42 BP; 13 A; 7 C; 7 G; 15 U; 0 other;

Query Match 100.0%; Score 29; DB 21; Length 42;
Best Local Similarity 100.0%; Pred. No. 0.00038;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 aaagauucuuuuuguaagccccaagggcu 29
||||||||||||||||||||||||||||||
Db 4 aaagauucuuuuuguaagccccaagggcu 32

RESULT 8
AAA71129
ID AAA71129 standard; RNA; 42 BP.
XX
AC AAA71129;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #198.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds -
XX

PS Example 7; Figure 126; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 42 BP; 13 A; 7 C; 7 G; 15 U; 0 other;

Query Match 100.0%; Score 29; DB 21; Length 42;
Best Local Similarity 100.0%; Pred. No. 0.00038;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 aaagauucuuuuuguaagccccaagggcu 29
||||||||||||||||||||||||||||||
Db 4 aaagauucuuuuuguaagccccaagggcu 32

RESULT 9
AAN60149
ID AAN60149 standard; cDNA; 825 BP.
XX
AC AAN60149;
XX
DT 13-SEP-1991 (first entry)
XX
DE Sequence encoding polypeptide murine interleukin-2 (IL-2) activity.
XX
KW T-cell biology; immune response; ss.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT CDS 49..558
FT /*tag= a
XX
PN EP177357-A.
XX
PD 09-APR-1986.
XX
PF 03-OCT-1985; 85EP-0307094.
XX
PR 05-OCT-1984; 84US-0658183.
PR 17-DEC-1986; 86US-0942600.
XX
PA (SCHE ) SCHERING BIOTECH CO.
PA (DNAX-) DNAX RES INST MOLEC.
XX
PI Lee FD, Yokota T, Arai K;
XX
DR WPI; 1986-095692/15.
DR P-PSDB; AAP60218.

```
XX
PT   New poly:peptide(s) showing murine interleukin-2 activity - are
PT   prepd. by recombinant DNA procedures with complementary DNA
PT   clones as vectors in escherichia coli
XX
PS   Claim 16; Fig 1; 53pp; English.
XX
CC   The pure murine IL-2 polypeptides facilitate the study of T-cell
CC   biology and the immune response, so that improvements in therapy are
CC   possible.
XX
SQ   Sequence 825 BP; 252 A; 179 C; 155 G; 239 T; 0 other;


  Query Match           100.0%;  Score 29;  DB 7;  Length 825;
  Best Local Similarity  69.0%;  Pred. No. 0.00064;
  Matches   20;  Conservative    9;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 aaagauucuuuuuguaagccccaagggcu 29
          |||||::|:::::|:||||||||||||||:
      688 aaagattcttttttgtaagccccaagggct 716


RESULT  10
AAN60906
ID   AAN60906 standard; cDNA; 940 BP.
XX
AC   AAN60906;
XX
DT   03-SEP-1991  (first entry)
XX
DE   Sequence encodes mouse interleukin-2 (MIL-2).
XX
KW   IL-2; lymphoma; ds.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   CDS             49..555
FT                   /*tag=  a
XX
PN   JP61058590-A.
XX
PD   25-MAR-1986.
XX
PF   29-AUG-1984;   84JP-0180086.
XX
PR   29-AUG-1984;   84JP-0180086.
XX
PA   (DAIL ) DAICEL CHEM IND KK.
XX
DR   WPI; 1986-117207/18.
DR   P-PSDB; AAP61032..
XX
PT   New imobilised microorganism - in porous particles of cellulose
PT   acetate.
XX
PS   Disclosure; Page 545; 13pp; Japanese.
XX
CC   Gne product may be expressed by a transformed host such as yeast, it
CC   is expressed in much larger quantities than from lymphoma and
CC   hybridoma cells.
CC   See also AAN60907-8.
XX
SQ   Sequence 940 BP; 288 A; 202 C; 173 G; 277 T; 0 other;


  Query Match           100.0%;  Score 29;  DB 7;  Length 940;
  Best Local Similarity  69.0%;  Pred. No. 0.00065;
  Matches   20;  Conservative    9;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 aaagauucuuuuuguaagccccaagggcu 29
          |||||::|:::::|:||||||||||||||:
Db    688 aaagattcttttttgtaagccccaagggct 716


RESULT  11
AAA70825
ID   AAA70825 standard; RNA; 45 BP.
XX
AC   AAA70825;
XX
DT   27-APR-2001  (first entry)
XX
DE   Molecular interaction site RNA #25.
XX
KW   Modulator; identification; molecular interaction; virtual library; ss.
XX
OS   Mus sp.
XX
PN   WO9958947-A2.
XX
PD   18-NOV-1999.
XX
PF   12-MAY-1999;   99WO-US10361.
XX
PR   12-MAY-1998;   98US-0076404.
PR   12-MAY-1998;   98US-0085092.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Ecker DJ,  Griffey R,  Crooke ST,  Sampath R,  Swayze E,  Mohan V;
PI   Hofstadler S,  McNeil J;
XX
DR   WPI; 2000-086439/07.
XX
PT   Identifying compounds which modulate activity of target biomolecules,
PT   used to provide compounds which can be used as pharmacological,
PT   agricultural and industrial compounds -
XX
PS   Claim 221; Page 232; 405pp; English.
XX
CC   This invention describes a novel method for identifying compounds which
CC   modulate the activity of a target biomolecule. The method uses
CC   3-dimensional representations of the biomolecule and a library of
CC   compounds and comprises (a) identifying at least one molecular
CC   interaction site of the target RNA; (b) generating in silico a virtual
CC   library of compounds predicted or calculated to interact with the
CC   molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC   representations of the target RNA with members of the virtual library of
CC   compounds to generate a hierarchy of the compounds ranked in accordance
CC   with their respective ability to form physical interactions with the
CC   molecular interaction site. The method also describes (1) RNA comprising
CC   a joined sequence of at least 24 nucleotides but not more than 70
CC   nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC   forming a first side of a first double stranded (ds) region; (b) 2
CC   nucleotides forming a first side of an internal loop region; (c) 4
CC   nucleotides forming a first side of a second ds region; (d) 4 or 5
CC   nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC   second side of the second ds region; (f) 4 nucleotides forming a second
CC   side of the internal loop region; and (g) 3 nucleotides forming a second
CC   side of the first ds region; (2) a purified and isolated RNA fragment
CC   comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC   methods and products can be used for identifying agents which modulate
CC   the activity of biomolecules, particularly RNA. Such agents can be used
CC   as pharmaceutical, agricultural or industrial compounds.
XX
SQ   Sequence 45 BP; 14 A; 7 C; 9 G; 15 U; 0 other;


  Query Match           96.6%;  Score 28;  DB 21;  Length 45;
  Best Local Similarity  100.0%;  Pred. No. 0.0011;
  Matches   28;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 aaagauucuuuuuguaagccccaagggc 28
          ||||||||||||||||||||||||||||
```

```
Db     18 aaagauucuuuuuguaagccccaagggc 45


RESULT 12
AAA70826
ID   AAA70826 standard; RNA; 45 BP.
XX
AC   AAA70826;
XX
DT   27-APR-2001  (first entry)
XX
DE   Molecular interaction site RNA #26.
XX
KW   Modulator; identification; molecular interaction; virtual library; ss.
XX
OS   Rattus sp.
XX
PN   WO9958947-A2.
XX
PD   18-NOV-1999.
XX
PF   12-MAY-1999;   99WO-US10361.
XX
PR   12-MAY-1998;   98US-0076404.
PR   12-MAY-1998;   98US-0085092.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Ecker DJ,  Griffey R,  Crooke ST,  Sampath R,  Swayze E,  Mohan V;
PI   Hofstadler S,  McNeil J;
XX
DR   WPI; 2000-086439/07.
XX
PT   Identifying compounds which modulate activity of target biomolecules,
PT   used to provide compounds which can be used as pharmacological,
PT   agricultural and industrial compounds -
XX
PS   Claim 222; Page 232; 405pp; English.
XX
CC   This invention describes a novel method for identifying compounds which
CC   modulate the activity of a target biomolecule. The method uses
CC   3-dimensional representations of the biomolecule and a library of
CC   compounds and comprises (a) identifying at least one molecular
CC   interaction site of the target RNA; (b) generating in silico a virtual
CC   library of compounds predicted or calculated to interact with the
CC   molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC   representations of the target RNA with members of the virtual library of
CC   compounds to generate a hierarchy of the compounds ranked in accordance
CC   with their respective ability to form physical interactions with the
CC   molecular interaction site. The method also describes (1) RNA comprising
CC   a joined sequence of at least 24 nucleotides but not more than 70
CC   nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC   forming a first side of a first double stranded (ds) region; (b) 2
CC   nucleotides forming a first side of an internal loop region; (c) 4
CC   nucleotides forming a first side of a second ds region; (d) 4 or 5
CC   nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC   second side of the second ds region; (f) 4 nucleotides forming a second
CC   side of the internal loop region; and (g) 3 nucleotides forming a second
CC   side of the first ds region; (2) a purified and isolated RNA fragment
CC   comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC   methods and products can be used for identifying agents which modulate
CC   the activity of biomolecules, particularly RNA. Such agents can be used
CC   as pharmaceutical, agricultural or industrial compounds.
XX
SQ   Sequence 45 BP; 14 A; 7 C; 9 G; 15 U; 0 other;


  Query Match             96.6%;  Score 28;  DB 21;  Length 45;
  Best Local Similarity   100.0%;  Pred. No. 0.0011;
  Matches   28;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 aaagauucuuuuuguaagccccaagggc 28
           ||||||||||||||||||||||||||||
```

```
Db     18 aaagauucuuuuuguaagccccaagggc 45


RESULT 13
AAA71088
ID   AAA71088 standard; DNA; 46 BP.
XX
AC   AAA71088;
XX
DT   27-APR-2001  (first entry)
XX
DE   Molecular interaction site DNA #111.
XX
KW   Modulator; identification; molecular interaction; virtual library; ss.
XX
OS   Unidentified.
XX
PN   WO9958947-A2.
XX
PD   18-NOV-1999.
XX
PF   12-MAY-1999;   99WO-US10361.
XX
PR   12-MAY-1998;   98US-0076404.
PR   12-MAY-1998;   98US-0085092.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Ecker DJ,  Griffey R,  Crooke ST,  Sampath R,  Swayze E,  Mohan V;
PI   Hofstadler S,  McNeil J;
XX
DR   WPI; 2000-086439/07.
XX
PT   Identifying compounds which modulate activity of target biomolecules,
PT   used to provide compounds which can be used as pharmacological,
PT   agricultural and industrial compounds -
XX
PS   Example 7; Figure 121; 405pp; English.
XX
CC   This invention describes a novel method for identifying compounds which
CC   modulate the activity of a target biomolecule. The method uses
CC   3-dimensional representations of the biomolecule and a library of
CC   compounds and comprises (a) identifying at least one molecular
CC   interaction site of the target RNA; (b) generating in silico a virtual
CC   library of compounds predicted or calculated to interact with the
CC   molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC   representations of the target RNA with members of the virtual library of
CC   compounds to generate a hierarchy of the compounds ranked in accordance
CC   with their respective ability to form physical interactions with the
CC   molecular interaction site. The method also describes (1) RNA comprising
CC   a joined sequence of at least 24 nucleotides but not more than 70
CC   nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC   forming a first side of a first double stranded (ds) region; (b) 2
CC   nucleotides forming a first side of an internal loop region; (c) 4
CC   nucleotides forming a first side of a second ds region; (d) 4 or 5
CC   nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC   second side of the second ds region; (f) 4 nucleotides forming a second
CC   side of the internal loop region; and (g) 3 nucleotides forming a second
CC   side of the first ds region; (2) a purified and isolated RNA fragment
CC   comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC   methods and products can be used for identifying agents which modulate
CC   the activity of biomolecules, particularly RNA. Such agents can be used
CC   as pharmaceutical, agricultural or industrial compounds.
XX
SQ   Sequence 46 BP; 14 A; 7 C; 9 G; 16 T; 0 other;


  Query Match             96.6%;  Score 28;  DB 21;  Length 46;
  Best Local Similarity   71.4%;  Pred. No. 0.0011;
  Matches   20;  Conservative    8;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 aaagauucuuuuuguaagccccaagggc 28
           ||||||:|:::::|:||||||||||||||
```

```
Db     19 aaagattctttttgtaagccccaagggc 46


RESULT 14
AAA71089
ID   AAA71089 standard; DNA; 46 BP.
XX
AC   AAA71089;
XX
DT   27-APR-2001  (first entry)
XX
DE   Molecular interaction site DNA #112.
XX
KW   Modulator; identification; molecular interaction; virtual library; ss.
XX
OS   Unidentified.
XX
PN   WO9958947-A2.
XX
PD   18-NOV-1999.
XX
PF   12-MAY-1999;   99WO-US10361.
XX
PR   12-MAY-1998;   98US-0076404.
PR   12-MAY-1998;   98US-0085092.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI   Hofstadler S, McNeil J;
XX
DR   WPI; 2000-086439/07.
XX
PT   Identifying compounds which modulate activity of target biomolecules,
PT   used to provide compounds which can be used as pharmacological,
PT   agricultural and industrial compounds -
XX
PS   Example 7; Figure 121; 405pp; English.
XX
CC   This invention describes a novel method for identifying compounds which
CC   modulate the activity of a target biomolecule. The method uses
CC   3-dimensional representations of the biomolecule and a library of
CC   compounds and comprises (a) identifying at least one molecular
CC   interaction site of the target RNA; (b) generating in silico a virtual
CC   library of compounds predicted or calculated to interact with the
CC   molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC   representations of the target RNA with members of the virtual library of
CC   compounds to generate a hierarchy of the compounds ranked in accordance
CC   with their respective ability to form physical interactions with the
CC   molecular interaction site. The method also describes (1) RNA comprising
CC   a joined sequence of at least 24 nucleotides but not more than 70
CC   nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC   forming a first side of a first double stranded (ds) region; (b) 2
CC   nucleotides forming a first side of an internal loop region; (c) 4
CC   nucleotides forming a first side of a second ds region; (d) 4 or 5
CC   nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC   second side of the second ds region; (f) 4 nucleotides forming a second
CC   side of the internal loop region; and (g) 3 nucleotides forming a second
CC   side of the first ds region; (2) a purified and isolated RNA fragment
CC   comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC   methods and products can be used for identifying agents which modulate
CC   the activity of biomolecules, particularly RNA. Such agents can be used
CC   as pharmaceutical, agricultural or industrial compounds.
XX
SQ   Sequence 46 BP; 14 A; 7 C; 9 G; 16 T; 0 other;


  Query Match             96.6%;  Score 28;  DB 21;  Length 46;
  Best Local Similarity   71.4%;  Pred. No. 0.0011;
  Matches   20;  Conservative    8;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 aaagauucuuuuuguaagccccaagggc 28
           ||||||::|:::::|:||||||||||||||
```

```
Db     19 aaagattctttttgtaagccccaagggc 46


RESULT 15
AAA71090
ID   AAA71090 standard; DNA; 46 BP.
XX
AC   AAA71090;
XX
DT   27-APR-2001  (first entry)
XX
DE   Molecular interaction site DNA #113.
XX
KW   Modulator; identification; molecular interaction; virtual library; ss.
XX
OS   Unidentified.
XX
PN   WO9958947-A2.
XX
PD   18-NOV-1999.
XX
PF   12-MAY-1999;   99WO-US10361.
XX
PR   12-MAY-1998;   98US-0076404.
PR   12-MAY-1998;   98US-0085092.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI   Hofstadler S, McNeil J;
XX
DR   WPI; 2000-086439/07.
XX
PT   Identifying compounds which modulate activity of target biomolecules,
PT   used to provide compounds which can be used as pharmacological,
PT   agricultural and industrial compounds -
XX
PS   Example 7; Figure 121; 405pp; English.
XX
CC   This invention describes a novel method for identifying compounds which
CC   modulate the activity of a target biomolecule. The method uses
CC   3-dimensional representations of the biomolecule and a library of
CC   compounds and comprises (a) identifying at least one molecular
CC   interaction site of the target RNA; (b) generating in silico a virtual
CC   library of compounds predicted or calculated to interact with the
CC   molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC   representations of the target RNA with members of the virtual library of
CC   compounds to generate a hierarchy of the compounds ranked in accordance
CC   with their respective ability to form physical interactions with the
CC   molecular interaction site. The method also describes (1) RNA comprising
CC   a joined sequence of at least 24 nucleotides but not more than 70
CC   nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC   forming a first side of a first double stranded (ds) region; (b) 2
CC   nucleotides forming a first side of an internal loop region; (c) 4
CC   nucleotides forming a first side of a second ds region; (d) 4 or 5
CC   nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC   second side of the second ds region; (f) 4 nucleotides forming a second
CC   side of the internal loop region; and (g) 3 nucleotides forming a second
CC   side of the first ds region; (2) a purified and isolated RNA fragment
CC   comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC   methods and products can be used for identifying agents which modulate
CC   the activity of biomolecules, particularly RNA. Such agents can be used
CC   as pharmaceutical, agricultural or industrial compounds.
XX
SQ   Sequence 46 BP; 14 A; 7 C; 9 G; 16 T; 0 other;


  Query Match             96.6%;  Score 28;  DB 21;  Length 46;
  Best Local Similarity   71.4%;  Pred. No. 0.0011;
  Matches   20;  Conservative    8;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 aaagauucuuuuuguaagccccaagggc 28
           ||||||::|:::::|:||||||||||||||
```

Db 19 aaagattctttttgtaagccccaagggc 46

Search completed: October 2, 2001, 05:03:57
Job time: 3809 sec

GenCore version 4.5


OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 05:01:49 ; Search time 57.41 Seconds
(without alignments)
95.628 Million cell updates/sec

Title: US-09-310-844B-25
Perfect score: 29
Sequence: 1 aaagauucuuuuuguaagccccaagggcu 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 94655562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| | Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|---|
| | 1 | 23.2 | 80.0 | 209 | 1 | US-08-510-032A-8 | Sequence 8, Appli |
| | 2 | 23.2 | 80.0 | 209 | 3 | US-08-688-514-8 | Sequence 8, Appli |
| | 3 | 23.2 | 80.0 | 801 | 6 | 5314995-8 | Patent No. 5314995 |
| c | 4 | 23.2 | 80.0 | 8491 | 2 | US-08-757-439-1 | Sequence 1, Appli |
| c | 5 | 17.6 | 60.7 | 72928 | 3 | US-09-009-913-1 | Sequence 1, Appli |
| | 6 | 17.4 | 60.0 | 51259 | 3 | US-08-781-891-209 | Sequence 209, App |
| c | 7 | 17.4 | 60.0 | 51259 | 3 | US-08-781-891-209 | Sequence 209, App |
| c | 8 | 17.2 | 59.3 | 2287 | 1 | US-08-222-619-1 | Sequence 1, Appli |
| c | 9 | 17.2 | 59.3 | 2287 | 4 | US-08-221-767-23 | Sequence 23, Appl |
| c | 10 | 17.2 | 59.3 | 2287 | 5 | PCT-US95-04075-1 | Sequence 1, Appli |
| c | 11 | 16.8 | 57.9 | 610 | 2 | US-08-928-926A-2 | Sequence 2, Appli |
| c | 12 | 16.8 | 57.9 | 610 | 3 | US-09-212-149-2 | Sequence 2, Appli |
| | 13 | 16.6 | 57.2 | 711 | 4 | US-08-998-416-900 | Sequence 900, App |
| c | 14 | 16.6 | 57.2 | 2446 | 3 | US-08-834-306-17 | Sequence 17, Appl |
| c | 15 | 16.6 | 57.2 | 2446 | 4 | US-08-993-674A-17 | Sequence 17, Appl |
| | 16 | 16.4 | 56.6 | 2311 | 2 | US-08-712-709-6 | Sequence 6, Appli |
| | 17 | 16.4 | 56.6 | 2311 | 3 | US-09-111-444-6 | Sequence 6, Appli |
| | 18 | 16.4 | 56.6 | 2311 | 4 | US-09-541-228-6 | Sequence 6, Appli |
| | 19 | 16.2 | 55.9 | 1201 | 3 | US-09-286-805-1 | Sequence 1, Appli |
| | 20 | 16.2 | 55.9 | 3383 | 1 | US-07-707-367-1 | Sequence 1, Appli |
| c | 21 | 16.2 | 55.9 | 3846 | 2 | US-08-845-161A-5 | Sequence 5, Appli |
| c | 22 | 16.2 | 55.9 | 3846 | 4 | US-09-270-751-5 | Sequence 5, Appli |
| c | 23 | 16.2 | 55.9 | 24417 | 2 | US-08-846-762-1 | Sequence 1, Appli |
| | 24 | 16.2 | 55.9 | 72928 | 3 | US-09-009-913-1 | Sequence 1, Appli |
| | 25 | 16 | 55.2 | 1803 | 3 | US-08-458-922-2 | Sequence 2, Appli |
| c | 26 | 16 | 55.2 | 6911 | 1 | US-08-311-174-4 | Sequence 4, Appli |
| | 27 | 15.8 | 54.5 | 325 | 2 | US-08-483-695-38 | Sequence 38, Appl |
| | 28 | 15.8 | 54.5 | 325 | 2 | US-07-965-285-38 | Sequence 38, Appl |
| | 29 | 15.8 | 54.5 | 325 | 2 | US-08-487-231-38 | Sequence 38, Appl |
| | 30 | 15.8 | 54.5 | 325 | 4 | US-09-201-912-38 | Sequence 38, Appl |
| c | 31 | 15.8 | 54.5 | 743 | 4 | US-08-821-994-38 | Sequence 38, Appl |
| c | 32 | 15.8 | 54.5 | 1102 | 4 | US-08-821-994-86 | Sequence 86, Appl |
| | 33 | 15.8 | 54.5 | 1207 | 1 | US-08-460-806-1 | Sequence 1, Appli |
| | 34 | 15.8 | 54.5 | 1207 | 1 | US-08-325-630-1 | Sequence 1, Appli |
| | 35 | 15.8 | 54.5 | 1333 | 4 | US-09-142-551A-1 | Sequence 1, Appli |
| c | 36 | 15.8 | 54.5 | 1390 | 4 | US-08-821-994-61 | Sequence 61, Appl |
| c | 37 | 15.8 | 54.5 | 1434 | 4 | US-08-821-994-62 | Sequence 62, Appl |
| c | 38 | 15.8 | 54.5 | 1441 | 4 | US-08-821-994-63 | Sequence 63, Appl |
| c | 39 | 15.8 | 54.5 | 2646 | 1 | US-08-365-189-7 | Sequence 7, Appli |
| c | 40 | 15.8 | 54.5 | 2688 | 1 | US-08-088-633-3 | Sequence 3, Appli |
| c | 41 | 15.8 | 54.5 | 2688 | 1 | US-08-245-756-3 | Sequence 3, Appli |
| c | 42 | 15.8 | 54.5 | 2688 | 1 | US-08-441-750-3 | Sequence 3, Appli |
| c | 43 | 15.8 | 54.5 | 2688 | 2 | US-08-441-751-3 | Sequence 3, Appli |
| c | 44 | 15.8 | 54.5 | 2688 | 5 | PCT-US92-02521-3 | Sequence 3, Appli |
| | 45 | 15.8 | 54.5 | 2696 | 1 | US-07-961-522-1 | Sequence 1, Appli |

ALIGNMENTS

```
RESULT   1
US-08-510-032A-8
; Sequence 8, Application US/08510032A
; Patent No. 5712126
;  GENERAL INFORMATION:
;    APPLICANT:  Sherman Weissman and Yarindra Prashar
;    TITLE OF INVENTION:  Analysis of Gene Expression By Display of 3'-
;    TITLE OF INVENTION:  end Restriction Fragments of cDNA
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Yahwak & Associates
;      STREET:  25 Skytop Drive
;      CITY:  Trumbull
;      STATE:  Connecticut
;      COUNTRY:  USA
;      ZIP:  06611
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  floppy disk
;      COMPUTER:  Macintosh
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  Microsoft Word 4.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/510,032A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  George M. Yahwak
;      REGISTRATION NUMBER:  26,824
;      REFERENCE/DOCKET NUMBER:  Yale
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (203)268-1951
;      TELEFAX:  (203)268-1951
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  209 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-08-510-032A-8


  Query Match             80.0%;  Score 23.2;  DB 1;  Length 209;
  Best Local Similarity   57.1%;  Pred. No. 0.051;
  Matches   16;  Conservative    9;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 aagauucuuuuuguaagccccaagggcu 29
           | ||:::|:::::|:|||||| | ||||:
Db      34 ATGATTCTTTTTGTAAGCCCTAGGGGCT 61
```

```
RESULT   2
US-08-688-514-8
; Sequence 8, Application US/08688514
; Patent No. 6010850
;  GENERAL INFORMATION:
;    APPLICANT:  Sherman Weissman and Yarindra Prashar
;    TITLE OF INVENTION:  Analysis of Gene Expression By Display of 3'-
;    TITLE OF INVENTION:  end Restriction Fragments of cDNA
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Yahwak & Associates
;      STREET:  25 Skytop Drive
;      CITY:  Trumbull
;      STATE:  Connecticut
;      COUNTRY:  USA
;      ZIP:  06611
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  floppy disk
;      COMPUTER:  Macintosh
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  Microsoft Word 4.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/688,514
;      FILING DATE:
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  George M. Yahwak
;      REGISTRATION NUMBER:  26,824
;      REFERENCE/DOCKET NUMBER:  Yale
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (203)268-1951
;      TELEFAX:  (203)268-1951
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  209 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-08-688-514-8


  Query Match           80.0%;  Score 23.2;  DB 3;  Length 209;
  Best Local Similarity   57.1%;  Pred. No. 0.051;
  Matches   16;  Conservative    9;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 aagauucuuuuuguaagccccaagggcu 29
           | ||::|:::::|:|||||| | ||||:
Db      34 ATGATTCTTTTTGTAAGCCCTAGGGGCT 61


RESULT   3
5314995-8
;Patent No. 5314995
;    APPLICANT: FELL, HENRY P.;GAYLE, MARGIT A.
;    TITLE OF INVENTION: THERAPEUTIC INTERLEUKIN-2-ANTIBODY
;BASED FUSION PROTEINS
;    NUMBER OF SEQUENCES: 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/468,390
;      FILING DATE: 22-JAN-1990
;SEQ ID NO:8:
;      LENGTH: 801
5314995-8


  Query Match           80.0%;  Score 23.2;  DB 6;  Length 801;
  Best Local Similarity   57.1%;  Pred. No. 0.063;
  Matches   16;  Conservative    9;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 aagauucuuuuuguaagccccaagggcu 29
           | ||::|:::::|:|||||| | ||||:
Db     651 atgattctttttgtaagccctaggggct 678


RESULT   4
US-08-757-439-1/c
; Sequence 1, Application US/08757439
; Patent No. 5866371
;  GENERAL INFORMATION:
;    APPLICANT:  BADZIONG, Werner
;    APPLICANT:  HABERMANN, Paul
;    APPLICANT:  MOELLER, Joerg
;    APPLICANT:  ARETZ, Werner
;    TITLE OF INVENTION:  PROCESS FOR USING THE YEAST ADH II
;    TITLE OF INVENTION:  PROMOTER SYSTEM FOR THE PRODUCTION OF HETEROLOGOUS
;    TITLE OF INVENTION:  PROTEINS IN HIGH YIELDS
;    NUMBER OF SEQUENCES:  1
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/757,439
;      FILING DATE:  27-NOV-1996
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE 19544233.4
;      FILING DATE:  28-NOV-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  SANDERCOCK, Colin G.
;      REGISTRATION NUMBER:  31,298
;      REFERENCE/DOCKET NUMBER:  18748/303/HOCE
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8491 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-757-439-1


  Query Match           80.0%;  Score 23.2;  DB 2;  Length 8491;
  Best Local Similarity   57.1%;  Pred. No. 0.094;
  Matches   16;  Conservative    9;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 aagauucuuuuuguaagccccaagggcu 29
           | ||::|:::::|:|||||| | ||||:
Db    6653 ATGATTCTTTTTGTAAGCCCTAGGGGCT 6626


RESULT   5
US-09-009-913-1/c
; Sequence 1, Application US/09009913
; Patent No. 6087485
;  GENERAL INFORMATION:
;    APPLICANT:  AxyS Pharmaceuticals, Inc.
;    TITLE OF INVENTION:  Asthma Related Genes
;    NUMBER OF SEQUENCES:  339
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Bozicevic & Reed, LLP
;      STREET:  285 Hamilton Ave, Suite 200
;      CITY:  Palo Alto
```

```
;       STATE:  CA
;       COUNTRY:  USA
;       ZIP:  94301
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette
;       COMPUTER:  IBM Compatible
;       OPERATING SYSTEM:  DOS
;       SOFTWARE:  FastSEQ for Windows Version 2.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/009,913
;       FILING DATE:  21-JAN-1998
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Sherwood, Pamela J
;       REGISTRATION NUMBER:  36,677
;       REFERENCE/DOCKET NUMBER:  SEQ-4P
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  650-327-3231
;       TELEFAX:  650-327-3231
;       TELEX:
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  72928 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  Genomic DNA
US-09-009-913-1


  Query Match             60.7%;  Score 17.6;  DB 3;  Length 72928;
  Best Local Similarity   45.8%;  Pred. No. 46;
  Matches   11;  Conservative    9;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       6 uucuuuuuguaagccccaagggcu 29
           :: ::::: :||| ||||| |||:
Db  37671 TTTTTTTTTTAAGTCCCAAAGGCT 37648


RESULT   6
US-08-781-891-209
; Sequence 209, Application US/08781891
; Patent No. 6090620
  GENERAL INFORMATION:
    APPLICANT:  Fu, Ying-Hui
    APPLICANT:  Yu, Chang-En
    APPLICANT:  Oshima, Junko
    APPLICANT:  Mulligan, John T.
    APPLICANT:  Schellenberg, Gerald D.
    TITLE OF INVENTION:  GENE AND GENE PRODUCTS RELATED TO
    TITLE OF INVENTION:  WERNER'S SYNDROME
    NUMBER OF SEQUENCES:  209
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  SEED and BERRY LLP
      STREET:  6300 Columbia Center, 701 Fifth Avenue
      CITY:  Seattle .
      STATE:  Washington
      COUNTRY:  USA
      ZIP:  98104-7092
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Floppy disk
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PatentIn Release #1.0, Version #1.30
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/781,891
      FILING DATE:  27-DEC-1996
      CLASSIFICATION:  800
    ATTORNEY/AGENT INFORMATION:
      NAME:  No. 6090620tenburg Ph.D., Carol
      REGISTRATION NUMBER:  39,317
      REFERENCE/DOCKET NUMBER:  240052.419
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  (206) 622-4900
      TELEFAX:  (206) 682-6031
  INFORMATION FOR SEQ ID NO:  209:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  51259 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  single
      TOPOLOGY:  linear
US-08-781-891-209


  Query Match             60.0%;  Score 17.4;  DB 3;  Length 51259;
  Best Local Similarity   48.1%;  Pred. No. 54;
  Matches   13;  Conservative    8;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       3 agauucuuuuuguaagccccaagggcu 29
           || ::|:::::| |||| |  || ||:
Db  26015 AGTTTCTTTTTGGAAGCACTTAGAGCT 26041


RESULT   7
US-08-781-891-209/c
; Sequence 209, Application US/08781891
; Patent No. 6090620
  GENERAL INFORMATION:
    APPLICANT:  Fu, Ying-Hui
    APPLICANT:  Yu, Chang-En
    APPLICANT:  Oshima, Junko
    APPLICANT:  Mulligan, John T.
    APPLICANT:  Schellenberg, Gerald D.
    TITLE OF INVENTION:  GENE AND GENE PRODUCTS RELATED TO
    TITLE OF INVENTION:  WERNER'S SYNDROME
    NUMBER OF SEQUENCES:  209
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  SEED and BERRY LLP
      STREET:  6300 Columbia Center, 701 Fifth Avenue
      CITY:  Seattle
      STATE:  Washington
      COUNTRY:  USA
      ZIP:  98104-7092
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Floppy disk
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PatentIn Release #1.0, Version #1.30
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/781,891
      FILING DATE:  27-DEC-1996
      CLASSIFICATION:  800
    ATTORNEY/AGENT INFORMATION:
      NAME:  No. 6090620tenburg Ph.D., Carol
      REGISTRATION NUMBER:  39,317
      REFERENCE/DOCKET NUMBER:  240052.419
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  (206) 622-4900
      TELEFAX:  (206) 682-6031
  INFORMATION FOR SEQ ID NO:  209:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  51259 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  single
      TOPOLOGY:  linear
US-08-781-891-209


  Query Match             60.0%;  Score 17.4;  DB 3;  Length 51259;
  Best Local Similarity   48.1%;  Pred. No. 54;
  Matches   13;  Conservative    8;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       3 agauucuuuuuguaagccccaagggcu 29
```

```
          || ::|:::::| |||| |  || ||:
Db  10496 AGTTTCTTTTTGGAAGCACTTAGAGCT 10470


RESULT   8
US-08-222-619-1/c
; Sequence 1, Application US/08222619
; Patent No. 5652352
;  GENERAL INFORMATION:
;    APPLICANT:  Lichenstein, Henri
;    APPLICANT:  Lyons, David
;    APPLICANT:  Wurfel, Mark
;    APPLICANT:  Wright, Samuel
;    TITLE OF INVENTION:  Afamin: A Human Serum Albumin-Like
;    TITLE OF INVENTION:  Protein
;    NUMBER OF SEQUENCES:  33
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amgen Center, Patent Operations/RRC
;      STREET:  1840 DeHavilland Drive
;      CITY:  Thousand Oaks
;      STATE:  California
;      COUNTRY:  U.S.
;      ZIP:  91320-1789
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/222,619
;      FILING DATE:
;      CLASSIFICATION:  435
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2287 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  318..2117
;    FEATURE:
;      NAME/KEY:  mat_peptide
;      LOCATION:  381..2114
;    FEATURE:
;      NAME/KEY:  sig_peptide
;      LOCATION:  318..380
  -08-222-619-1


  Query Match             59.3%;  Score 17.2;  DB 1;  Length 2287;
  Best Local Similarity   54.5%;  Pred. No. 40;
  Matches   12;  Conservative    7;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      3 agauucuuuuuguaagccccaa 24
          ||| :|:::::|: ||| ||||
Db    175 AGACTCTTTTTGTGAGCACCAA 154


RESULT   9
US-08-221-767-23/c
; Sequence 23, Application US/08221767
; Patent No. 6268212
;  GENERAL INFORMATION:
;    APPLICANT:  Simonet, William S.
;    APPLICANT:  Lichenstein, Henri S.
;    APPLICANT:  Lyons, David E.
;    TITLE OF INVENTION:  Tissue Specific Transgene Expression
;    NUMBER OF SEQUENCES:  24
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amgen Inc., U.S. Patent Operations/NAO
;      STREET:  1840 Dehavilland Drive
;      CITY:  Thousand Oaks
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  91320-1789
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/221,767
;      FILING DATE:
;      CLASSIFICATION:  800
;  INFORMATION FOR SEQ ID NO:  23:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2287 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  318..2117
;    FEATURE:
;      NAME/KEY:  mat_peptide
;      LOCATION:  381..2114
;    FEATURE:
;      NAME/KEY:  sig_peptide
;      LOCATION:  318..380
US-08-221-767-23


  Query Match             59.3%;  Score 17.2;  DB 4;  Length 2287;
  Best Local Similarity   54.5%;  Pred. No. 40;
  Matches   12;  Conservative    7;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      3 agauucuuuuuguaagccccaa 24
          ||| :|:::::|: ||| ||||
Db    175 AGACTCTTTTTGTGAGCACCAA 154


RESULT  10
PCT-US95-04075-1/c
; Sequence 1, Application PC/TUS9504075
;  GENERAL INFORMATION:
;    APPLICANT:  AMGEN INC.
;    TITLE OF INVENTION:  Afamin: A Human Serum Albumin-Like
;    TITLE OF INVENTION:  Protein
;    NUMBER OF SEQUENCES:  33
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amgen Center, Patent Operations/RRC
;      STREET:  1840 DeHavilland Drive
;      CITY:  Thousand Oaks
;      STATE:  California
;      COUNTRY:  U.S.
;      ZIP:  91320-1789
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/04075
;      FILING DATE:
;      CLASSIFICATION:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2287 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  cDNA
;    FEATURE:
```

```
;       NAME/KEY:  CDS
;       LOCATION:  318..2117
;     FEATURE:
;       NAME/KEY:  mat_peptide
;       LOCATION:  381..2114
;     FEATURE:
;       NAME/KEY:  sig_peptide
;       LOCATION:  318..380
PCT-US95-04075-1


  Query Match             59.3%;  Score 17.2;  DB 5;  Length 2287;
  Best Local Similarity   54.5%;  Pred. No. 40;
  Matches   12;  Conservative    7;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 agauucuuuuuguaagccccaa 24
           ||| :|:::::|: ||| ||||
Db     175 AGACTCTTTTTGTGAGCACCAA 154


RESULT  11
US-08-928-926A-2/c
; Sequence 2, Application US/08928926A
; Patent No. 5925543
;  GENERAL INFORMATION:
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Tang, Y. Tom
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  NADH DEHYDROGENASE B17 SUBUNIT
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/928,926A
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0384 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  610 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  TBLYNOT01
;      CLONE:  44898
US-08-928-926A-2


  Query Match             57.9%;  Score 16.8;  DB 2;  Length 610;
  Best Local Similarity   50.0%;  Pred. No. 48;
  Matches   14;  Conservative    7;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 aaagauucuuuuuguaagccccaagggc 28
           | | | :|:::::|:|  ||||| || |
Db     318 AGATACTCTTTTTGTATACCCCATGGAC 291


RESULT  12
US-09-212-149-2/c
; Sequence 2, Application US/09212149
; Patent No. 6100036
;  GENERAL INFORMATION:
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Tang, Y. Tom
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  NADH DEHYDROGENASE B17 SUBUNIT
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/212,149
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/928,926
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0384 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  610 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  TBLYNOT01
;      CLONE:  44898
US-09-212-149-2


  Query Match             57.9%;  Score 16.8;  DB 3;  Length 610;
  Best Local Similarity   50.0%;  Pred. No. 48;
  Matches   14;  Conservative    7;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 aaagauucuuuuuguaagccccaagggc 28
           | | | :|:::::|:|  ||||| || |
Db     318 AGATACTCTTTTTGTATACCCCATGGAC 291


RESULT  13
US-08-998-416-900
; Sequence 900, Application US/08998416
; Patent No. 6239264
;  GENERAL INFORMATION:
;    APPLICANT:  Philippsen, Peter
;    APPLICANT:  Pohlmann, Rainer
;    APPLICANT:  Steiner, Sabine
;    APPLICANT:  Mohr, Christine
;    APPLICANT:  Wendland, Jurgen
```

```
;    APPLICANT:  Knechtle, Philipp
;    APPLICANT:  Rebischung, Corinne
;    TITLE OF INVENTION:  GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
;    TITLE OF INVENTION:  AND USES THEREOF
;    NUMBER OF SEQUENCES:  1152
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6239264artis Corporation
;      STREET:  3054 Cornwallis Road
;      CITY:  Research Triangle Park
;      STATE:  No. 6239264th Carolina
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/998,416
;      FILING DATE:  24-DEC-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  CH 0016/97
;      FILING DATE:  31-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  PF/5-30306/A/CGC1976
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-541-8587
;      TELEFAX:  919-541-8689
;  INFORMATION FOR SEQ ID NO:  900:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  711 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;      ORGANISM:  PAG1561UP
US-08-998-416-900


  Query Match             57.2%;  Score 16.6;  DB 4;  Length 711;
  Best Local Similarity   47.8%;  Pred. No. 61;
  Matches   11;  Conservative    8;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       6 uucuuuuuguaagccccaagggc 28
           ::|:::::|:   | ||||||||
       168 TTCTTTTTGTTTTCACCAAGGGC 190


RESULT  14
US-08-834-306-17/c
; Sequence 17, Application US/08834306
; Patent No. 6054135
;  GENERAL INFORMATION:
;    APPLICANT:  Reed,Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Lodes, Michael J.
;    APPLICANT:  Houghton, Raymond L.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR THE DETECTION AND   PREVENTION OF T
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,306
;      FILING DATE:  15-APR-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.422C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2446 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-834-306-17


  Query Match             57.2%;  Score 16.6;  DB 3;  Length 2446;
  Best Local Similarity   56.5%;  Pred. No. 75;
  Matches   13;  Conservative    6;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       6 uucuuuuuguaagccccaagggc 28
           ::|:: : |:|  ||||||||||
Db     370 TTCTTCTCGTAGTCCCCAAGGGC 348


RESULT  15
US-08-993-674A-17/c
; Sequence 17, Application US/08993674A
; Patent No. 6228372
;  GENERAL INFORMATION:
;    APPLICANT:  Reed,Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Lodes, Michael J.
;    APPLICANT:  Houghton, Raymond L.
;    APPLICANT:  Smith, John M.
;    APPLICANT:  McNeill, Patricia D.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR THE DETECTION AND   PREVENTION OF
;    NUMBER OF SEQUENCES:  81
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/993,674A
;      FILING DATE:  18-DEC-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.422C2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2446 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
```

US-08-993-674A-17

Query Match 57.2%; Score 16.6; DB 4; Length 2446;
Best Local Similarity 56.5%; Pred. No. 75;
Matches 13; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

```
Qy      6 uucuuuuuguaagccccaagggc 28
          ::|:: : |:|  ||||||||||
Db    370 TTCTTCTCGTAGTCCCCAAGGGC 348
```

Search completed: October 2, 2001, 05:01:52
Job time: 5319 sec